OFFICINA
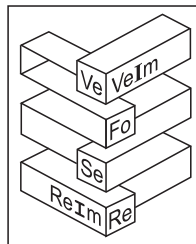## TEXTOLOGICA
18.

Fifteen years
of research
in text linguistics

The Officina Textologica project

edited by

István Csűry

Officina Textologica 18.

Edited by
István Csűry

Reviewed by
Péter B. Furkó

Layout
Beáta Bagaméry

Cover design
József Varga

Contents

1.
Debrecen studies on text and discourse:
plural approaches to a single object
On the path of János S. Petőfi[1]
ISTVÁN CSŰRY

1. Introduction

The present volume provides the non-Hungarian speaking community of linguists interested in text and discourse studies with an overview of the research carried out around the periodical *Officina Textologica*, published by the Institute of Hungarian Linguistics, University of Debrecen, Hungary, since 1997. Authors try to offer some insights into publications available so far in Hungarian only, breaking this way the limits of diffusion imposed by the fact that *Officina Textologica* has only collected articles written in this somewhat rare language.

In the past sixteen years, annual workshops preceding the elaboration of the forthcoming issues have gathered scholars not only from the departments of languages and linguistics of this university but also from other Hungarian universities as well as from abroad. Their fields of interest, theoretical and methodical approaches are often quite different; however, a coherent and fruitful dialogue is established each time on the grounds of a unique theoretical framework called *semiotic textology*.

This volume is not about semiotic textology *in general*. Readers can find several publications on this topic in languages other than Hungarian as well, among which Giuffrè (2011) is a recent and exhaustive account of the theory. What will be dealt with on the following pages is *a polyglot research program conceived in a semiotic-textological framework*.

Semiotic textology and the name of János Sándor Petőfi are inseparable. Not only was he the founder of the theory that has been evolving since the early 1970s and the inventor of its successive designations but he stood behind almost every research project referring to it. It was the case in Hungary where he played an essential role in boosting research into text and discourse especially after the end of the communist era. The *Officina Textologica* project too grew from his inspiration and from his intensive cooperation with the Debrecen team. Not surprisingly, this volume was to be introduced by Petőfi himself with a brief overview of the state of the art. However, the article has never been written.

Editorial works were about to be finalized and we were waiting only for Petőfi's article when the master, as many of us considered this fragile man of an exceptional brightness and intellectual energy, passed away. For all who had the chance of collaborating with him, it has been an irreparable loss.

The original purpose of this volume was not a commemoration for János Sándor Petőfi. Nevertheless, we very much owe it to him to dedicate it to his memory. Let us therefore begin by recalling some essential facts about his life and oeuvre. This will also give us a better idea about the context of the *Officina Textologica* project.

2. From theory building to research organization

János Sándor Petőfi was born in Miskolc on 23th April 1931. His career can be divided into seven main periods, four of which (the first three and the last one) are related to Hungary while the intermediary ones to Sweden, Western Germany, and Italy, respectively. He graduated at Kossuth Lajos University (today's University of Debrecen) where he first studied mathematics, physics and descriptive geometry (until 1955) and, later, German language and literature (until 1961). He worked as a secondary school teacher in Debrecen and Budapest till 1964 when he was nominated research fellow at the Centre for Computer Science at the Institute of Linguistics of the Hungarian Academy of Sciences (HAS). For five years, he had Ferenc Kiefer and György Szépe, his contemporaries for co-workers, who, just like himself, became decisive figures of modern Hungarian linguistics. It was the latter who encouraged Petőfi to take the road that led him to semiotic textology – and, in a concrete, geographical sense, far from his country for a long time.

He pursued his scientific activities at the University of Umeå (Sweden, 1969–1970) and at the University of Konstanz (Germany, 1971). After having obtained his PhD degree and habilitation title in Umeå (1971), he received professorship and was nominated chair of the Department of Semantics of the Faculty of Linguistics and Literary Theory at the University of Bielefeld (1972).

The next stage of his itinerary, after seventeen years spent in Germany, was the University of Macerata, in Italy, where he held the professorship of philosophy of language from 1989 till his retirement. Although retired in 2007, he was nominated director of the *Centro di Documentazione e ricerca sugli approcci semiotico-testologici alla multi ed intermedialità* being just created, in capacity of professor emeritus. Petőfi led this research centre in collaboration with his disciples. He was discharged of his duties in 2011 on his own request. In the meantime, his presence in Hungary became more and more intense after the regime change. He considered 2007 as the beginning of the seventh period of his career.

The first document of his oeuvre in linguistics is his university degree thesis on a German language history topic (Umschreibungen in dem Werk "Der Arme Heinrich" von Hartmann von Aue. Grammatische und stilistische Untersuchungen). Soon after that, he revealed the real breadth of his endeavour, which spanned literary science, linguistics, philosophy of language, semiotics and communication theory. He was familiar with the works written in Russian as well as in English but it never led him to simply adopt the ideas of the Soviet or American linguists: even his very first papers attest an original conception of how texts should be analyzed. His PhD thesis, published as a monograph under the title *Transformationsgrammatiken und eine ko-textuelle Texttheorie*, was widely acclaimed and brought him international fame. Thenceforth, his works played a decisive role in the development of textology (known as text linguistics and text grammar as well).

One can find no discontinuity in his oeuvre even if his conception of text has evolved. In fact, there are two main periods in the development of his theory. The first one embraces the 1970s when he focused on the modelling of the internal – or, using his own term, cotextual – factors of textuality while the theory itself was elaborated during the second period. Its starting point is the *text structure – world structure* (TeSWeST) theory that he later renamed as *semiotic textology* and to which he made continuous improvements. In this framework, text is considered as a complex sign with regard to its function in communication, having, despite of its usual verbal prevalence, multimediality for an essential property. Thus, it is from a semiotic point of view that one might account for every (i.e. syntactic, semantic and pragmatic) aspect of its interpretation. Conceived in this way, text is far from being an exclusive object of linguistics, as its description necessitates an interdisciplinary approach. Thus, presenting the disciplinary context of text research was always an essential part of his work. Interdisciplinarity meant for Petőfi the respect of the plurality of scientific knowledge, an effective method of problem solving as well as a way of living and thinking rather than a simple slogan. It is for that very reason that the theoretical framework of semiotic textology could become a common platform of exchanges for many scholars with various orientations, as it is precisely the case of our polyglot research program in textology at the University of Debrecen. A synthesis of the theory written in Hungarian was published in 2004 by Akadémiai Kiadó publishing house under the title of *A szöveg mint komplex jel* (*Text as a Complex Sign*).

Petőfi was an extremely productive author. His bibliography contains hundreds of publications written in seven languages, published in renowned international journals and book series. Among the titles, one can find about two dozen monographs (signed either as a single author or as a co-author), and three dozen collective volumes appeared under his editorship. He promoted teaching

of text linguistics in Hungary as well by contributing to textbooks. He was founding editor of two book series that belong to the most noted forums in the field of text research (*Research in Text Theory*, de Gruyter [1977], *Papiere zur Textlinguistik*, Buske [1974]). He also founded a review with Teun A. van Dijk (*Text*), and even two Hungarian periodicals have been created with his collaboration: *Szemiotikai Szövegtan* (*Semiotic Textology*, 1990, with Imre Békési and László Vass) in Szeged and *Officina Textologica* (1997, with Irma Szikszainé Nagy and Edit Dobi) in Debrecen.

János S. Petőfi is one of the best-known Hungarian linguists in the world. This charismatic scholar was not only a pioneer of his field (or, better said, fields) but a decisive personality in organizing scientific life whom a multitude of talented linguists in many places all over the world acknowledges as their master. He spent more than half of his working years abroad, however, he never lost contact with his country and with Hungarian scholars; on the contrary, he provided them with all the help he could offer. After the regime change, he played an increasingly salient role in the activities of Hungarian textological research groups inside and outside the borders of the country, at the universities of Pécs, Debrecen, Szeged, Budapest, and Cluj. If their research could be institutionalized and linked to international circuits, it is also due to the merit of Petőfi who endeavoured to align scientific activities carried out in the context of different cultures.

All the communities, from the smallest to the largest ones, that he distinguished by the support of his (not only) spiritual presence have been keen not to seem ungrateful. As a matter of fact, János S. Petőfi was a widely acclaimed scholar. He attended three Nobel symposia as an invited participant, and even a Nobel Prize ceremony as an honorary guest. (Nobel Symposia, held since 1965, gather world-class researchers from areas of science where breakthroughs are occurring or deal with topics of primary cultural or social significance.) He was distinguished by the János Lotz medal of the International Association for Hungarian Studies in 1996. He was honorary member of the Hungarian Association for Semiotic Studies as well as of the Hungarian Association of Applied Linguists and Language Teachers (HAALLT). He was conferred the Doctor Honoris Causa award by three different universities (two in Hungary: Pécs [1991], Debrecen [1996] and the University of Turin [2004]) and the *magister emeritus* award by Gyula Juhász Teachers' Training College (Szeged). The Hungarian Academy of Sciences elected him an external member in 2007. As for himself, he undoubtedly saw the greatest honour in the accomplishment of his work by his disciples and in their results. Their gratitude and their attachment clearly manifested themselves by the fact that the participants, representing Hungarian, Italian and German colleagues, at the conference held at HAS on 27th April 2011 offered him not less than four

festschriften on his 80th birthday

3.   A systematic exploration of an uneven domain

When Petőfi returned to his alma mater with the idea of the *Officina Textologica* project, there was no research group in text linguistics he could directly address. But there were several linguists, attached to different departments of languages (Hungarian, English, French and German), working on different topics in various fields, with different theoretical backgrounds, who shared a common interest in problems of text and discourse. The interdisciplinary nature of semiotic textology that Petőfi proposed to adopt as an overall framework managed to yield an excellent platform of dialogue. Taking also into account the centre of gravity these departments formed with regard to external cooperation possibilities, the future participants considered it obvious to set up an organized form of scientific exchange following an organically conceived long-term program. It was on these grounds that the keynote paper of Petőfi was published in 1977 as the first volume of the *Officina Textologica* series. The topics dealt with are the following:
1. The disciplinary environment of text study. Text linguistics and textology in text study.
2. The relation of these terms and research orientations in the relevant literature.
3. Aspects of the text linguistic/textological analysis and description of text creating factors.
4. Linguistics from a textological perspective.
5. Semiotic textology as a theoretical framework for text linguistics.

In the last chapter, Petőfi sketches the structure and the main thematic groups of the planned publications on the basis of his introductory considerations.

The *Officina Textologica* project was originally conceived as a series of annual meetings on specific topics, followed each time by the publication of a collective volume, with the objective of creating "a special forum (or fictive roundtable) for scholars working on and interested in the problems of text study" (Petőfi 1997: 84). This might include either the construction of a general theory of texts, studies on a given language from a textual point of view, or the creation of textological/text linguistic tools that are applicable in a specific field. The authors, indeed, remained in their particular fields, maintaining their specific point of views while observing the same object in a textological approach. However, the framework defined by Petőfi ensured the coherence of the dialogue and, thus, that of the thematic volumes. The discussion was intended to be polyglot and integrative as much as possible, i.e. multilingual/contrastive studies were planned and a kind of paradigm was to be set up in order to allow

combining and/or explicitly comparing perspectives of researchers from different background and interests.

This project, having been supported by the Hungarian Scientific Research Fund (OTKA) during two funding periods, has yielded so far two cycles of thematic volumes. The first one, including fourteen numbers published between 1997 and 2008, follows Petőfi's original program. After his above mentioned, introductory study comes a volume on coreference containing analyses of Hungarian texts (Petőfi 1998). It is followed by another collection of papers of the same kind on problems of possible linear arrangements of sentence constituents (Szikszainé Nagy 1999). The fourth volume is a discussion of coreference relations presented in the second one (Dobi and Petőfi 2000). The following issue contains studies on a wider range of topics, pertaining to the relationship of grammar, text linguistics and textology (Petőfi and Szikszainé Nagy 2001). The next year's volume revisits the problems of linear arrangement in a form of discussion (Szikszainé Nagy 2002). Volume 7 introduces contrastive text linguistics with papers on linearization and theme/rheme structure (Petőfi and Szikszainé Nagy 2002). It is followed by Edit Dobi's monograph on a two-step representation of text sentences in a semiotic-textological framework (Dobi 2002). In 2003, the authors follow the exploration of contrastive text linguistics, addressing this time another aspect of linearization, i.e. thematic progression (Petőfi and Szikszainé Nagy 2003). Volume 10 is about the textual role of conceptual schemata (Petőfi and Szikszainé Nagy 2004) while in volume 11 three authors discuss textological works written in Hungarian (Petőfi 2005). Volume 12 deals with the question of co-referentiality using a contrastive approach (Petőfi and Szikszainé Nagy 2005). Another monograph follows as volume 13, written by myself on connectives (Csűry 2005). Finally, the last number of this first period takes up the problems of conceptual organization by examining the role of scenarios in building texts (Dobi 2008).

The second cycle begins in 2009, with another programmatic volume of Petőfi (Petőfi 2009). He gives first a critical and detailed overview of the work carried out so far in the framework of the *Officina Textologica* project and establishes a positive balance with regard to the aims set at the beginning of the project. After several chapters on coherence and its approaches, he proposes that the team focus their attention on the study of coherence in texts and presents the planned topics of the next five volumes. According to this plan, the following issue (published as one of the four Festschriften offered to Petőfi on his 80th birthday in 2011) provides an evaluative overview of the terminology used in the (multilingual i.e. Hungarian, English, French and German) literature on particular areas of text linguistics (Dobi 2011). Volume 17 deals with phenomena relating to textual meaning on the basis of contrastive analyses of

semantic organization on English, German and French corpora (Dobi 2012).

4. Five aspects of text as a complex sign

In the present volume, seven authors, having regularly participated in the *Officina Textologica* project, review the essential results of these publications in five main fields that have been most intensively explored so far. They also intend to point out some specific problems and questions that have been left open, thus illustrating not only their particular preoccupations but also the possible (and probable) research orientations of *Officina Textologica*.

After summarizing the studies related to the problem of meaning at the level of discourse published in the volumes of *Officina Textologica*, Péter Pelyvás's paper on *Meaning at the level of discourse: from lexical networks to conceptual frames and scenarios* presents the cognitive framework in which concepts spanning from lexical networks to conceptual frames and scenarios can be useful for textual research.

In their study, Andrea Nagy and Franciska Skutta deal with *Co-reference*. They first present the essential concepts and the coreference model of János S. Petőfi. The second part of their work turns briefly to the twelve articles of volume 2 of the series in order to treat some special questions raised by the types of texts analyzed in this volume. Finally, they outline further research on the subject, as it appears in two, more recent volumes.

Edit Dobi's contribution is a paper on problems of linearization and information structure (*On the Results of the Discussion about the Phenomenon of Linearization of Text Sentences*). Linearization is an issue for both sentence grammar and text linguistics; therefore, one is confronted whith the question where the border lies between these two fields. The author aims to provide a review of the studies published in the *Officina Textologica* that are relevant to this topic, and presents not only the findings of the authors but further issues to discuss and problems to solve as well.

István Csűry's study (*Connectives and discourse markers*) deals with describing structural and pragmatic markers in the framework of textology. After reviewing *Officina Textologica* publications devoted to connectives and discourse markers, he discusses the main problems of identifying and classifying such elements and proposes a simple yet complete and useful way to tell apart text/discourse structuring element types. Terminological issues are also addressed in this chapter, and analyses of several text excerpts are presented as well in order to illustrate the interface role of connectives between syntactic, informational and discourse structures.

Finally, Károly István Boda and Judit Porkoláb present *Semiotic-textological approaches to literary discourse*. Their main concern is to overview the basic

methods and formalism of co-reference analysis, developed from the first volume of *Officina Textologica*. It has become a powerful tool to explore the textological structure and thematic composition of literary texts. The paper contains a table as well indicating the literary texts of which one can find (partial or comprehensive) analyses in the *Officina Textologica* volumes.

References

Csűry, István 2005. *Kis könyv a konnektorokról*. Officina Textologica Vol. 13. [Small book on connectives]. Debrecen: Debreceni Egyetem, Magyar Nyelvtudományi Tanszék.

Dobi, Edit 2002. *Kétlépcsős szövegmondat-reprezentáció szemiotikai textológiai keretben*. Officina Textologica Vol. 8. [A Two-Step Representation of Text Sentences in a Semiotic-Textological Framework].

Dobi, Edit, ed. 2008. *A forgatókönyv mint dinamikus szövegszervező erő*. Officina Textologica Vol. 14. [Scripts as dynamic text organisers]. Debrecen: Debreceni Egyetem, Magyar Nyelvtudományi Tanszék.

Dobi, Edit, ed. 2011. *A szövegösszefüggés elméleti és gyakorlati megközelítési módjai. poliglott terminológiai és fogalmi áttekintés*. Officina Textologica Vol. 16. [Theoretical and empirical approaches to coherence. A polyglott overview of terminology and concepts]. Debrecen: Debreceni Egyetemi Kiadó.

Dobi, Edit, ed. 2012. *A szövegösszefüggés elméleti és gyakorlati megközelítési módjai. Diszkusszió*. Officina Textologica Vol. 17. [Theoretical and empirical approaches to coherence. Discussion]. Debrecen: Debreceni Egyetemi Kiadó.

Dobi, Edit and S. János Petőfi, eds. 2000. *Koreferáló elemek - koreferenciarelációk. Magyar nyelvű szövegek elemzése. Diszkusszió*. Officina Textologica Vol. 4. [Coreferential elements, co-reference relations. Analysis of Hungarian language texts. Discussion]. Debrecen: Kossuth Egyetemi Kiadó.

Giuffrè, Mauro 2011. Introduction. In: Sprachtheorie *und germanistische Linguistik, Supplement 1*. Münster: Nodus Publikationen. 1-7.

Petőfi, S. János 1997. *Egy poliglott szövegnyelvészeti-szövegtani kutatóprogram*. Officina Textologica Vol. 1. [A polyglot research program in textology/text linguistics]. Debrecen: Kossuth Egyetemi Kiadó.

Petőfi, S. János 2009. *Egy poliglott szövegnyelvészeti-szövegtani kutatóprogram II. Adalékok a verbális szövegek szövegösszefüggőség-hordozóinak vizsgálatához*. Officina Textologica Vol. 15. [A polyglot research program in textology/text linguistics II. Contributions to the study of coherence markers of verbal texts]. Debrecen: Debreceni Egyetem Magyar Nyelvtudományi Tanszék.

Petőfi, S. János, ed. 1998. *Koreferáló elemek - koreferenciarelációk. Magyar nyelvű szövegek elemzése*. Officina Textologica Vol. 2. [Coreferential elements, co-reference relations. Analysis of Hungarian language texts]. Debrecen: Kossuth Egyetemi Kiadó.

Petőfi, S. János, ed. 2005. *Adalékok a magyar nyelvészet szövegtani diszkurzusához. Három közelítés*. Officina Textologica Vol. 11. [Contributions to the textological discourse in Hungarian linguistics. Three approaches]. Debrecen: Debreceni Egyetem, Magyar Nyelvtudományi Tanszék.

Petőfi, S. János and Irma Szikszainé Nagy, eds. 2001. *Grammatika - szövegnyelvészet - szövegtan*. Officina Textologica Vol. 5. [Grammar - text linguistics - textology]. Debrecen: Kossuth Egyetemi Kiadó.

Petőfi, S. János and Irma Szikszainé Nagy, eds. 2002. *A kontrasztív szövegnyelvészet aspektusai. Linearizáció: téma–réma szerkezet*. Officina Textologica Vol. 7. [Aspects of contrastive text linguistics. Linearization: theme/rheme structure]. Debrecen: Kossuth Egyetemi Kiadó.

Petőfi, S. János and Irma Szikszainé Nagy, eds. 2003. *A kontrasztív szövegnyelvészet aspektusai. Linearizáció: tematikus progresszió*. Officina Textologica Vol. 9. [Aspects of contrastive text linguistics. Linearization: thematic progression]. Debrecen: Kossuth Egyetemi Kiadó.

Petőfi, S. János and Irma Szikszainé Nagy, eds. 2004. *A szövegorganizáció elemzésének aspektusai. Fogalmi sémák*. Officina Textologica Vol. 10. [Aspects of analysis of text organization. Conceptual schemata]. Debrecen: Debreceni Egyetem, Magyar Nyelvtudományi Tanszék.

Petőfi, S. János and Irma Szikszainé Nagy, eds. 2005. *A korreferencialitás poliglott vizsgálata*. Officina Textologica Vol. 12. [Polyglot study of coreference]. Debrecen: Debreceni Egyetem, Magyar Nyelvtudományi Tanszék.

Szikszainé Nagy, Irma, ed. 1999. *Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. Magyar nyelvű szövegek elemzéséhez*. Officina Textologica Vol. 3. [Towards the analysis of the linear arrangement of text sentence constituents. Analysis of Hungarian language texts.]. Debrecen: Kossuth Egyetemi Kiadó.

Szikszainé Nagy, Irma, ed. 2002. *Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. Magyar Nyelvű szövegek elemzéséhez. Diszkusszió*. Officina Textologica Vol. 6. [Towards the analysis of the linear arrangement of text sentence constituents. Analysis of Hungarian language texts. Discussion]. Debrecen: Kossuth Egyetemi Kiadó.

2.
Meaning at the level of discourse: from lexical networks to conceptual frames and scenarios[1]
PÉTER PELYVÁS

Summaries of studies related to the problem of meaning at the level of discourse published in the volumes of Officina Textologica

Officina Textologica has devoted volumes to the analysis of meaning at levels higher than the clause. Vol. 10 is a collection of papers on conceptual schemes, Vol. 14 is devoted to scenarios. The summaries of the papers included in these volumes are given below. Throughout the volumes of the series, a number of papers are (at least partially) concerned with meaning at different levels of representation: in argument structure, in the organization of the tense-aspectual frame of a clause, in coreference relationships, etc. We cannot undertake to discuss them all here.

After the summaries, we propose to give a more or less consistent sample of how an originally sentence-oriented theory, holistic cognitive grammar is capable of bridging the traditional gap between sentence linguistics and text linguistics, by applying methods originally proposed for describing larger units to the analysis of a number of factors that are essential in the organization of the clause. This is based on some of the papers by Péter Pelyvás.

Officina Textologica 10, Aspects of the analysis of the organization of texts: conceptual schemas

Conceptual schemas play a central role in the analysis of the compositional organisation of texts. The thorough exploration of its various aspects was the core subject of a thematic conference held at the University of Debrecen on December 10th, 2004, the presentations of which are included in volume 10 of Officina Textologica.

In *'Various aspects of the analysis of the relations providing context'*, JÁNOS S. PETŐFI directs attention to the representation of constringency (i.e. the verbal manifestations of the real or assumed relationship between facts, see 1.3) as a fundamental aspect of the analysis of the context. Among the various relations

---

providing context, the author points out the relations between microcompositional units of text and conceptual schemas. In this respect, he deals with a special thesauristic representation of cognitive frames.

In her study *'Cognitive frames, reference, pronouns'*, ANDREA CSŰRY gives a representation of the role of certain indefinite pronouns of the French language by way of performing a detailed analysis of four text segments. Her considerations are based on conceptual schemas, that is, cognitive frames and scripts.

In *'The role of cognitive frames in poetic texts'*, KÁROLY I BODA and JUDIT PORKOLÁB elaborate a specific cognitive model for the interpretation process of poems. Their approach to the interpretation process is based on the selection of appropriate concordances from various sources which can be linked to the poem to be interpreted. The corpus, which is the source of the concordances, forms a computer-based world of texts. Its hypertextual organisation leads to a specific model for the interpretation process where the examination of cognitive frames plays a central role.

In her study *'Conceptual frames and context in the short story «Omlette à Woburn» by Dezső Kosztolányi'*, ÁGNES DE BIE KERÉKGYÁRTÓ gives a cognitive analysis of the short story. The central concept of the author's theory is that the successful interpretation of a text — that is, the text-based process of its meaning — is based on the harmonised mobilisation of the writer's and reader's knowledge of the world.

In his study *'WRITING as a specific cognitive objectivation'*, LÁSZLÓ JAGUSZTIN discusses the different aspects of the relationship between writing (or text) and the world as it is reflected in the short story "Kinevez... Tetik hadnagy" by Tinyanov.

In *'Filling in indefinite places'*, FRANCISKA SKUTTA interprets a few introductory paragraph from the novel *"A gyertyák csonkig égnek"* by Sándor Márai. The author concentrates on features of the context that can only be interpreted with recourse to information that is based — beyond the verbally expressed context — on the reader's own knowledge of the world. Her final conclusion is that the "indefinite places" (Roman Ingarden) may never be filled in entirely.

In his study *'Ways of decoding'*, SÁNDOR KISS analyses the first chapter of the novel *"Fanni hagyományai"* by József Kármán in order to illucidate the decoding process of interpretation during which the reader's knowledge of the world of text develops. In order for this process to be successful, the author attributes a special role to the knowledge of the cognitive frames that can be attached to text to be interpreted.

In *'How to create strange vocabularies?'*, ISTVÁN CSŰRY deals with the representation possibilities of cognitive frames and scripts. The author takes

standard lexicological practice as the starting point of his considerations in order to raise theoretical and practical issues concerning a thesaurus which can serve as a representation of conceptual schemas. As for the problems of describing conceptual schemas, the author analyses selected examples to illustrate the problems that arise while describing conceptual schemas.

In *'Analysing and ways of formalising cognitive frames in specialised texts'*, EDIT DOBI and ÁKOS KUKI try to reveal the role and significance of the formal description in the characterisation of the semantic relations occurring between the elements of the cognitive frames that can be attached to the same text. Analysing a relatively simple part of a specialised text as well as the cognitive frames that cover it, the authors try to explore and formalise the structure of the semantic relations between the elements of the cognitive frames which reflect the semantic structure of the analysed text.

Officina Textologica 14, The scenario as a dynamic force in organizing texts

This volume of *Officina Textologica* deals with the (partly or completely) semantic aspects of *context*. Following previous volumes which dealt with co-referential relationships, thematic progression and (cognitive) frames respectively, this volume contains selected essays on *scripts*[2].

The essays are written versions of the presentations held at the conference *Scripts as dynamic text organisers* in Spring 2007. We might also well add ".. first approximation" to the title since, as is usual with *Officina Textologica,* further detailed and in-depth discussion of the topic will follow in a subsequent volume.

The assumptions that the authors elaborate in this varied and colourful volume emanate from different theoretical backgrounds and views. There may, for instance, be substantial differences in how the authors define and interpret the basic concept of *script*. They may consider a *script* as

● specific parts of background knowledge that belong to the collective knowledge of a community, or (in a perhaps slightly more individual interpretation),

● a level of subjective knowledge that assumes some specialised knowledge regarding e.g. the creation process of a literary work, a poet's course of life, etc.

KÁROLY ISTVÁN BODA and JUDIT PORKOLÁB adopt two different approaches to the concept of *script*. In a textological framework, they try to explore an interpretation of the concept of *script* that could be appropriate in the

---

[2] Some contributions use the term *script*, others use *scenario* to talk about essentially the same concept. My personal preference is for the latter due to its extended use in the cognitive literature. But I will leave other authors' choices unchanged and regard the two terms as synonymous in this discussion.

communication, text processing and understanding process. Within a *cognitive science* framework, the interpretation of the concept of *script* is based on the background knowledge that can be arranged in a script-like form. As a consequence, it is necessary to examine different types of knowledge first. The authors describe four types of knowledge, along with the types of scripts that can be associated with them. This approach provides a broad interdisciplinary framework for research on the use of textological methods in the representation of cognitive process.

In her essay EDIT DOBI examines the possible relationship(s) between the type of text and the type(s) and organisation of the script(s) which are to be explored in the text. In general, two conclusions of the research can be outlined: first, the analyses indicate that promising and well applicable results can be foreseen in the field of textology and text typology. Second, the results depend crucially on the way the concept of *script* is defined, that is, how the degree of complexity of its constituents is established. For example, we may assume that one possible script for the event of "arrival at a restaurant" is as follows: we enter, look for a table, take off our coat, sit down etc. (with some concessions regarding relative order). At some point, we have to decide whether this script provides satisfactory detail of description or we must take into consideration specific scripts concerning the way *how* we take off our coats, the various rituals of sitting down at the table etc. Beyond these issues, in the summary of the essay further questions are formulated for the future research of *scripts*.

In his essay, ISTVÁN CSŰRY discusses some basic theoretical and practical questions of script research. The author evaluates, among others, the significance or "linguistic/textological usefulness" of the study of scripts either on the macro level (i.e. in the whole text) or the micro level (i.e. in specific parts of text). In the analysis of scripts, he finds it important to pay special attention to *connectives,* which can be characteristic of certain organisations of scripts. In order to demonstrate his ideas on scripts, he examines the place and function of connectives in dialogues.

Distinguishing between the language-related and real-world aspects of scripts, SÁNDOR KISS outlines the phenomenon of the so-called "*shifting script*". Shifting scripts are defined as "modified patterns" which describe a "modified course of events". The author's approach to the concept of *script* is basically traditional but can also be characterised as innovative in a sense: he refines the classical interpretation of the *script* by emphasizing the fact that there can be more than one linguistic realisations of a script describing a typical course of events. The author characterises the concept of "*shifting scripts*" by the use of the four rhetorical operations (addition, deletion, substitution, and rearrangement). In order to illustrate his ideas, the author gives colourful literary examples from short stories by *Iván Mándy*.

ANDREA CSŰRY studies the scripts of dialogues and, similarly to Sándor Kiss, concentrates on those characteristics that are different from accepted prototypes. While analysing dialogues, she intends to reveal and illustrate the process of *misunderstanding*. Relying on Roman Jacobson's model of communication, the author examines all aspects of communication that, as possible sources of errors, can lead to misunderstanding. These aspects are as follows: linguistic and non-linguistic knowledge of the sender and receiver, the message, code, medium and context. The varied and vivid sample texts, which come from both everyday life and literature, all serve the author's intention to give instructive models for the process of misunderstanding which is basically stereotyped but can nevertheless have a number of interesting variations.

In her essay, FRANCISKA SKUTTA examines the relationship between two remarkable and complex phenomena: she investigates the related elements of, and differences between *script* and *synopsis* with the aim of exploring connections between them. The comparison is facilitated by the fact that both can be considered as systems (i.e. sets of organised elements). After outlining an elaborate typology of synopses the author focuses on the *narrative synopsis,* the study of which is most helpful in exploring relationships between script and synopsis. She establishes that one evident similarity between scripts and synopses is as follows: events and participants in both of them are "*beyond time*" and exist "*in themselves*", i.e. they are in a "*timeless present*" and do not have the "*narrator's contribution*". The two phenomena can be seen as being even more closely related: the author demonstrates a kind of mutual dependence between script and narrative synopsis, which leads to the conclusion that "textological and narrative research can both provide major contributions to the other's scientific enrichment".

ANNAMÁRIA KABÁN interprets the concept of *script* in a way which reminds one of the applied sciences. She considers *scripts* basically as dynamic plans or strategies of organisation underlying the construction of texts. To demonstrate her ideas, she analyses the poem *Psalmus Hungaricus* by *Jenő Dsida*. In the interpretation process she emphasizes a special function of *scripts* which activates, as a loosest script, certain regions of the interpreter's background knowledge concerning the history of literature. Therefore she considers some crucial elements of background knowledge related to the interpretation process– e.g. the religious faith of the poet, Psalm 137, which provides a frame of genre for the interpretation process, rhetorical devices, etc.–as *scripts*. As a final conclusion she proposes that the overall script of Dsida's poem consists in "how the refusal of values becomes a value".

In his essay BÉLA LÉVAI also relies on a literary work as a framework for analysing the concept of *script*. While examining the poem *Favágó* (Woodman) by *Attila József* in Hungarian and in its Russian translation, he focuses on the

writing process of the poem and adopts Gábor Tolcsvai-Nagy's definition of the *script*. He compares the original Hungarian poem with its Russian translation regarding the appearance and organisation of *poetical script*, and finds substantial differences. It is very interesting for the reader to follow how the original script of the poet can be recognised in, or interpreted into, the Russian translation. The differences come mainly from the characteristics of the two languages.

As it was mentioned before, the analyses and interpretations of the concept of *script* in the essays of this volume of *Officina Textologica* emanate from more or less different theoretical backgrounds. As a result, the conclusions and questions of the authors and the results of their research provide various suggestions for future directions of script research or, more generally, for the investigation of the semantic organisation of text.

Finally, a highly relevant paper from a regular author of *Officina Textologica*, which was published in a different collection:

CSŰRY I. 2011. A forgatókönyv mint elméleti kategória és kommunikációs eseménytípus multimodális megközelítésben. [The script as a theoretical category and as a type of communicative event in a multimodal framework.] In: Enikő Németh T. (ed.): *Ember-gép kapcsolat. A multimodális ember-gép kommunikáció modellezésének alapjai.* Budapest: Tinta Könyvkiadó. 145− 178.

From lexical networks to conceptual frames and scenarios: the cognitive framework

1. Characteristics of the cognitive framework

1.1. Generative grammar and the traditional linguistic paradigm

Since many of the *Officina* papers discussed in this section are part of an endeavour to apply Langacker's holistic cognitive grammar to the analysis of structures beyond the clause/sentence level, it is natural to begin our discussion with a brief introduction to the principles and methods of this approach to language and its use.

Cognitive grammar differs significantly from traditional approaches to text in that its interest in structures larger than clauses or sentences develops organically from its psychologically based holistic view of all phenomena connected with language and its use – already at the lowest levels of organization. The system was originally developed in the 1980's with an aim to overcome at least some of the difficulties and contradictions inherent in traditional sentence grammars (especially Chomsky's Generative Grammar and truth-functional semantics) but it was soon realized by its founders (Lakoff and Johnson 1980, Langacker 1987, 1991) that this could only be achieved by breaking away from almost all the

tenets of the Saussurian and Chomskyan tradition that had been at the foundation of a system-based modular approach to grammar. This tradition emphasized predictability and compositionality at all levels of linguistic description by stating that the task of linguistics was to account for the ideal native speaker's ability to create and understand novel sentences on the basis of an autonomous system of rules that were clearly separable from general processes of human cognition to the extent that they had to be presumed to be innate.

The most obvious objection to the generative system in the 1980's was that, in order to achieve full predictability of grammatical phenomena, it had to continually impose severe limitations on what was to be regarded as part of grammar (originally formulated by Chomsky (1964: 62) as observational adequacy: '*the lowest level, indicating whether the grammar has properly identified the phenomena that need to be accounted for'.* In addition to the distinction of *competence* vs. *performance*, already present in the Saussurian tradition, this led to the dichotomies of *grammar* vs. *lexicon*, *core grammar* vs. *periphery*, *UG principles* vs. *parsing rules* at various stages in the development of Chomskyan theory, all with the net effect of reducing the scope of grammar and, as Newmeyer (1991) claims for the last distinction, a separation of innate linguistic knowledge from non-innate general conversational (parsing) principles. This is a special point of interest in our discussion here since it creates an enormous gap between the language system and its use for communication – ultimately between sentence grammar and text linguistics.

Formal semantics (in its weakest interpretation) is the application of the rules of formal logic to meaning in natural language (to the extent that that is possible). There are a number of objections even to this weak interpretation that space does not allow us to discuss in detail here. I would only like to emphasize that a combination of the generative interpretation of linguistic competence (defined as the ability to create and understand novel sentences) with its strict separation from any non-linguistic knowledge must naturally lead to the rule of full compositionality that is also inherent in formal semantics. After all, if novel sentences are *not* understood relying only on the meanings of the component parts and their syntactic arrangement, what other factors could be involved? On the other hand, the question arises of how much of actual language remains semantically analyzable if the rule of strict compositionality is retained? Is there a difference in terms of compositionality between (1a) and (1b)? If there is one, how can it be accounted for?

> (1) a Mary has a chocolate in her mouth.
>   b Mary has a cigarette in her mouth.

1.2. The cognitive alternative

Owing at least partially to these considerations, the most important point of departure of a cognitive alternative has had to be a break away from system linguistics, formal semantics and the rule of compositionality. We do not have the space here to give anything like a thorough introduction to Cognitive Grammar, we will only concentrate on some of its basic assumptions (based on Langacker 1987, 1991) that are most relevant to our purposes in this paper.

● Cognitive Grammar is psychologically rather than logically based. It defines language as a means of cognition as well as communication, claiming that the system bears every mark of having been elaborated for use for both purposes by humans. As a result, it is a usage-based approach that does not make a distinction between linguistic competence and performance on the one hand, or between linguistic and non-linguistic knowledge on the other.

● As a result, it does not need to rely on the principle of strict compositionality. The meaning of complex structures (or *units,* in the cognitive terminology) is only motivated by the meanings of the component parts and the way they are assembled, additional information comes from the general (and often varying) cognitive background of language users. It is true that the grammar loses some predictive power in this way, but as we have referred to it in Section 1.1, this power seems to have been a burden rather an asset to generative grammar as well, forcing it to continually restrict its professional interest to structures that do not resist their kind of analysis. Cognitive grammar, on the other hand, is capable of accounting for the (strictly semantic, communicative or social) motivation of the structures that are actually used, making predictions as to what other structures might or might not be used for the semantic purposes on hand.

● Cognitive grammar denies the direct reflection of logical relationships in grammatical structure (often referred to as *logical-grammatical relationships*). A discrepancy between the logical and the grammatical form is the sign of a transformation for the generativist. Cognitive grammar does not admit transformations, holding the view that different grammatical structures result from different conceptualizations. A key issue in this approach to language is the notion of construal.

● Construal gives the language user considerable freedom in deciding the question '*What is going on?*' when a set of events needs to be conceptualized. Different construals are (often only slightly but sometimes radically) different conceptualizations of a situation, which will in turn lead to different linguistic forms at all levels of organization beginning with lexical networks (argument structure) to questions related to the organization of discourse.

The generativist and the cognitive approaches to relationships of meaning and form are compared in Figure 1.
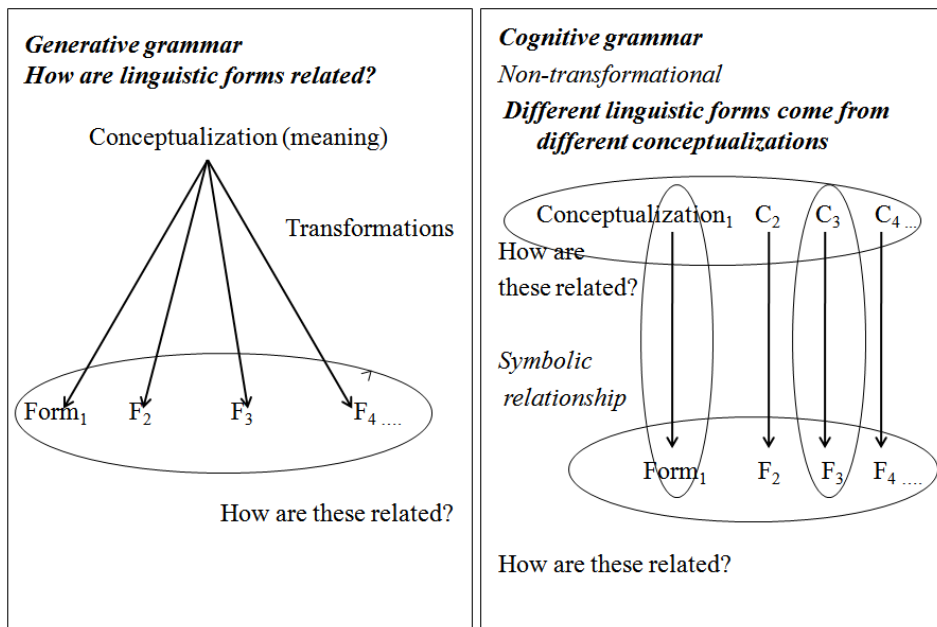
Figure 1: Generative and cognitive grammar

The psychological process of construal is essential in the organization of the cognitive framework. The key notions of *scope* (deciding what is in profile, what is essential or marginal in a conceptualization), *prominence* (the primary distinction of figure and ground and a secondary one within the figure) and *perspective* (the degree of speaker involvement: objective vs. subjective viewing arrangement) are all based on construal, and they in their turn are determining factors in the grammatical organization of language structures at all levels.

The secondary distinction of *trajector* and *landmark* within the figure, for instance, determine subject and object selection: a crucial factor in organizing a clause. This view of grammatical functions can also explain why purely semantic definitions of subject and object have always failed in linguistics: the determining factor is *attention* (tr/lm selection) and semantic factors may (or may not) have only an indirect influence on this choice.

- A related factor that has a very important role in the shaping of grammatical form is the formation of Idealized Cognitive Models (ICMs). The conceptualizer, in assessing a situation, is not given ready-made solutions. With an active effort, (s)he has to make some sense of what is going on or form an ICM: *a situation, its participants and the relationships that hold among them, as construed by the conceptualizer* (Lakoff 1987).

In summary: Over the years, attempts have been made to apply the methods of sentence linguistics to texts--with little success, owing to the inefficiency in this field of the tools it was able to use. Holistic cognitive grammar, based on the language user's assessment of a situation (ICM) relying on a full knowledge of the world available to him/her from all possible sources, seems capable of bringing sentence linguistics and text linguistics closer together because it already analyses sentences with tools designed for the analysis of larger contexts or scenarios. In the following sections I will give examples of how this could work, beginning with the relevance of alternative cognitive construals in argument structure, through the significance of ICMs in communication and in the construal of scenarios, and concluding with a brief cognitive analysis of epistemic grounding (modality), a process that anchors what is said to the knowledge of the speaker and the hearer about the world.

2. Attempts at cognitive solutions: lexical networks and conceptual frames

2.1. Argument structure: load

A simple case of the choices involved in the formation of an ICM is the selection of an image schema, but that selection will determine argument structure in the clause, as in the case of the English word *load* (Pelyvás 2001 in *Officina Textologica* 5, an English version can be found in Pelyvás 1996).

Pairs of sentences like (2a) and (b) have been something of a problem for modern theories of language ever since Fillmore (1977) brought them into the focus of attention:

> (2) a John loaded hay onto the truck.
>   b John loaded the truck with hay.

Early generative grammar attempted to analyze the pair as transformationally related, but the attempt had to be given up partly because no transformational mechanism could be found or created to link them (especially in GB) and partly because there is an obvious difference in the meaning of the two. Since cognitive grammar holds the view that different (but related) forms come from different (but related) conceptualizations, our task is now to find out what these conceptualizations are and how they are related.

The first thing to notice is that the event described (which may be, 'objectively' speaking, 'identical' in the two sentences), can be divided into two subevents or subtrajectories, since they both involve motion:

1. John's physical activity (prototypically a repeated movement of the arms [tools] along a well-defined trajectory). This part is identical for both *a* and *b*:

2. The subtrajectories 'observed' or conceptualized here are already different:

    for *a*: the hay changed location

for *b*: a container was filled

As for subtrajectory 2, it could be argued that both events have to occur in both sentences: you cannot fill a container with hay without the hay changing location. Objectively speaking, that may be true, but cognitive grammar has the remarkable characteristic of allowing for the conceptualizer's ability to structure reality in different ways:

A fundamental notion of cognitive semantics is that a predication does not reside in conceptual content alone but necessarily incorporates a particular way of construing and portraying that content. Our capacity to construe the same content in alternate ways is referred to as **imagery**; expressions describing the same conceived situation may nonetheless be semantically quite distinct by virtue of the contrasting images they impose on it. (Langacker 1991: 4)

Owing to the difference in the construal of subtrajectory 2, the selection of landmark is changed: *the hay* is in profile in *a* and *the truck* in *b*. The expected consequence is the change in argument structure. The landmark becomes the direct object in active sentences[3].

There is substantial evidence from grammar that we have the schema of a container in (2b), which is not present in (2a):

● *(2b) is telic, (2a) is atelic.* One of the prototypical properties of a container is that it has a certain capacity or volume and when that volume is filled, the process cannot go on. This corresponds to the requirement that a telic process must have a natural conclusion.

Note that (2a) could only be 'made telic' by limiting the amount of hay available. The simplest way to do this is by using a definite NP:

(3) John loaded the hay onto the truck,

but the rather atypical case of filling a definite volume with exactly the amount of substance that is available is perhaps less than fully acceptable.

---

[3] Note that in Fillmore's Case Grammar (Fillmore 1968)*the truck* always had to be locative, partly because there was no separate case for a *container* and partly because of *the objective view taken of the situation*. Preserving deep case relationships was essential during whatever transformations the sentence underwent. In cognitive grammar, since the situation is construed subjectively, there is nothing to prevent the speaker from regarding *the truck* as a container in one case and simply as location in the other.

(4) ?John loaded the truck with the hay (3.8)[4]

● *The ICM of filling a container has some constraints on the substance used.* Gradual, or, in the case of solids, repeated action is typically involved. The substance used must fill the whole volume of the container, so it must have the properties typically expressed by a mass noun or plural count noun. Compare:

(5) John loaded the truck with
	hay	(5.0)
	peas	(5.0)
	bricks	(5.0)
	machines	(4.1)
	*a car	(2.4)

None of these NPs would be problematic at all with the structure in (2a).

● *The criteria for filling a container properly and for moving or transporting hay are not exactly the same.* Compare:

(6) John did not load the truck properly:
   a. a lot of hay was left in the field   (3.6)
   b. it was left half empty   (3.9)
   c. he was certain to lose half of
        the hay on his way home   (3.4)

(7) John did not load the hay on the truck properly:
   a. a lot of hay was left in the field   (2.9)
   b. it was left half empty   (3.1)
   c. he was certain to lose half of
        the hay on his way home   (4.5)

The scores here are not always really definitive, but seem to support our argument.

In this case study my aim has been to show that construal in terms of *imagery* (whether or not to apply the **container** image schema to *truck*) has direct consequences on the argument structure of sentence pairs like (2a) and (b).

2.2. Argument structure: correction of an ICM

Sometimes it may be necessary for a speaker to discard a cognitive model seen as appropriate for describing a situation at the time of observation in favour of another one seen now as more adequate. This is typically an issue that would

---

[4] The numbers in brackets against this and some of the examples to follow are grades of acceptability (1 to 5) based on a survey of a small group of native and non-native speakers of English.

never come up as such in a system grammar, but the grammatical consequences of such a move would need to be dealt with in a systematic way. Unfortunately, this is very often not the case in traditional grammars, where the transformationally related alternatives given in (8) were clearly treated as synonymous in the 1970's and even more recent developments such as the rule-to-rule hypothesis only state that every syntactic rule has some counterpart in semantics, without feeling the need to examine the nature of the semantic difference.

As we have seen, cognitive grammar changes the relationship of the components arguing that it is changes in conceptualization that have syntactic consequences rather than the other way round. The case of *load* was a relatively simple one. The sentences in (8), traditionally seen as structurally related by the transformation of Raising or by Exceptional Case Marking are of greater complexity (Pelyvás 2001 in *Officina Textologica* 5, for a full English version see Pelyvás 2011b):

> (8) a I saw *Steve steal your car,* but at the time I thought that he was only borrowing it.
> b I saw *Steve stealing your car,* but …
> c *I saw *that Steve stole your car*, but …

In order to understand why the Raising construction is a suitable tool for the purpose, we have to look into the cognitive theory of epistemic grounding. In terms of Langacker (1991) Tense and Modality (which, according to Pelyvás (1996, 2011a,b) can also be expressed by cognitive predicates like *see* or *think/believe*) serve as grounding predications that relate an event to the circumstances of its utterance: speaker/hearer knowledge, time and other deictic elements. It can be hypothesized that the non-finite form occurring in the subordinate clause of the construction, with its less-than-fully grounded status, is in a symbolic relationship with this conceptual content of correction.

The difference between (8a) and (b) on the one hand and (8c) on the other is not in the *grounding* of the whole structure (something that the speaker does at the time of speaking) but in that of the subordinate structure marked in italics. The less than fully grounded non-finite form indicates a (now corrected) problem in conceptualization or ICM formation (*borrowing* vs. *stealing*), something that the conceptualizer does (or rather did) at the time of perception. The event was *not* conceptualized as stealing.

To find further support and also a higher level of generalization for the hypothesis that the forms appearing in the complement of a cognitive predicate are in a symbolic relationship with its status relative to grounding, we can also examine Hungarian. This language almost totally lacks Raising but still seems to

have a much wider array of choices in the expression of ICM correction. Consider the possible Hungarian equivalents of the English sentences in (8):

(9) a Láttam,   *hogy* Pista   **ellopta*   az autódat,
       I-see-Past *that* Steve   *steal-Perf.-Past*   your car
         de akkor azt hittem,   hogy csak kölcsönveszi.
         but then that I-believe-Past   that only he-borrow-Pres.
      = relative past
     b   *?ellopja*
           steal-Perf. Present = relative tense
     c   **lopta*
           steal-Imperf. Past
     d   **lopja*
           steal-Imperf. Present = relative tense

   The unacceptable (9a) combines a finite object clause with Past Tense which is to be seen here as *absolute:* it relates the time of the situation to the time of utterance, giving it fully grounded status, in contrast to the *relative tense* appearing in (9b). The Present Tense form of (9b) relates the time of the event 'only' to the time of the matrix clause, but even that change will make the sentence only marginally acceptable. The imperfect forms in (9 c and d) only make the situation worse: they appear to strengthen a false link between seeing something and conceptualizing it as stealing at the time of the event.

   In (10) the object clause is replaced with a clause of manner, which improves the situation considerably, since the sentence is now *more* about the ingredients of the ICM that were observable to the conceptualizer at the time of conceptualization than about his/her formation of an (incorrect) cognitive model.

(10) a Láttam,   *ahogy* Pista *ellopta*   az autódat,
        I-see-Past   *how*   Steve *steal-Perf.-Past* your car
          de   akkor azt   hittem …
          but   then that   I-believe-Past …
      b   *ellopja*
                            steal-Perf. Present = relative tense

   In (11) we have a time clause in subordination, which only permits absolute tense. The marginal acceptability of (11b) may be attributable to the fact that the imperfect form, in opposition to its role in (9), an object clause, now marks the incompleteness of the experience, making its conceptualization more difficult. This contrast is similar to the difference between the English sentences in (8a) and (8b):

(11) a Láttam, *amikor* Pista *ellopta* az autódat,
de akkor azt hittem …
I-see-Past *when* Steve *steal-Perf.-Past* your car
but then that I-believe-Past …

b ?*lopta*
steal-Imperf. Past

Finally, structures similar to English Raising are also possible in Hungarian, even though only (12a) would be more than a very rough equivalent. In (12b) to (12d) the subject NP is easily seen as part of the conceptual content of the matrix clause as well:

(12) a Láttam *Pistát* *ellopni* az autódat, de akkor
azt hittem …
I-see-Past *Steve-Acc* *steal-Inf.* your car but then
…

b Láttam *Pistát,* *ahogy ellopta*
I-see-Past *Steve-Acc* *as/how he-steal-Past*

az autódat, de akkor azt hittem …
your car but then that I-believe-Past …

c *ahogy* *ellopja*
as/how *he-steal-Present = relative tense*

d *amikor* *ellopta*
when *he-steal-Past*

The aim of this Section has been to illustrate on the examples of English and Hungarian how alternative argument structures seen as (often meaningless) transformations in traditional grammar can express subtle differences in the speaker's attitude to what (s)he has to say. Grammatical differences reflect differences in the creation or correction of Idealized Cognitive Models. In Section 3 we will see an example of how different ICMs of the same situation in different people's minds can affect communication.

2.3. Tense and Aspect

At a higher level of discourse, it can be shown that the construal of scenarios (both in the sense of apprehending an event and of relating it in conversation) are very consistently reflected in grammatical structure. The Simple Past Tense may be sufficient to relate a set of events 'as they happened'. But humans have a strong tendency to highlight anteriority or simultaneity relations or cause–effect

relationships, etc. as well. This requires more sophisticated grammatical tools even at the level of sentence structure (progressive and perfect forms or passives – universally seen as a means of expressing 'marked' topic–comment relations. (The issue is discussed in detail in Pelyvás 2008, in *Officina Textologica* 14.) For a quick illustration, consider the encounter described in (13), the key to an exercise for students of English:

> (13) The very moment I saw the man I **found** him suspicious. I **suspected seeing** him somewhere before, but where? **Could** he **be** the man I **had been cheated by** shamelessly just a few weeks before? Before I **had answered** that question, it suddenly **dawned** on me that he **must be** the clerk I **had been having** affairs with at the bank for some time. He **had been said to have been arrested** for some kind of serious offence on the job, but now, obviously, it **could not have been** true. He pretended **not to have noticed** me, just as I **had decided** I **would do** myself, which **saved** me a lot of trouble. In fact I **had been hoping** he **would do** exactly that. I **do not know** what I **would have done** if he **had decided to come** up to me and **shake** hands. I **am** certainly **not looking forward to meeting** him again, and I`d rather he never **showed up** in the future, if he **could** help it, either!

As native speakers would probably agree, the story *could* be told in simpler terms as well. But something of the message would certainly be lost.

2.4. Coreference: conceptual structure in deontic modality

As we have remarked in Section 1, Cognitive Grammar was from the beginnings deeply dissatisfied with the application of formal logic to the description of meaning (cf. motivatedness vs. compositionality). One of the areas where a clear alternative may offer itself is the description of modality[5]. The first significant step was Sweetser (1990), a work that suggests that

● modals are to be described in terms of force dynamics;

● the epistemic meanings of the modals are the result of metaphorical extension from their root (prototypically deontic) senses.

Sweetser (1990) provides a very simple conceptual schema consisting only of forces and barriers, which, although a good point of departure, can be shown to be erroneous in a number of ways (cf. Pelyvás 1996). In subsequent work (e.g. Pelyvás 2011b) Pelyvás suggests a more sensitive analysis in which barriers are replaced by counteracting forces to account for the potentiality and flexibility in

---

[5] This discussion is largely based on based on Pelyvás (2005) in Officina Textologica 12. A full English version is available in Pelyvás (2011a).

the modal system and also, more important for our discussion here, that these forces are to be associated by the participants of the situation (an important step in the creation of an ICM, cf. 1.2.).
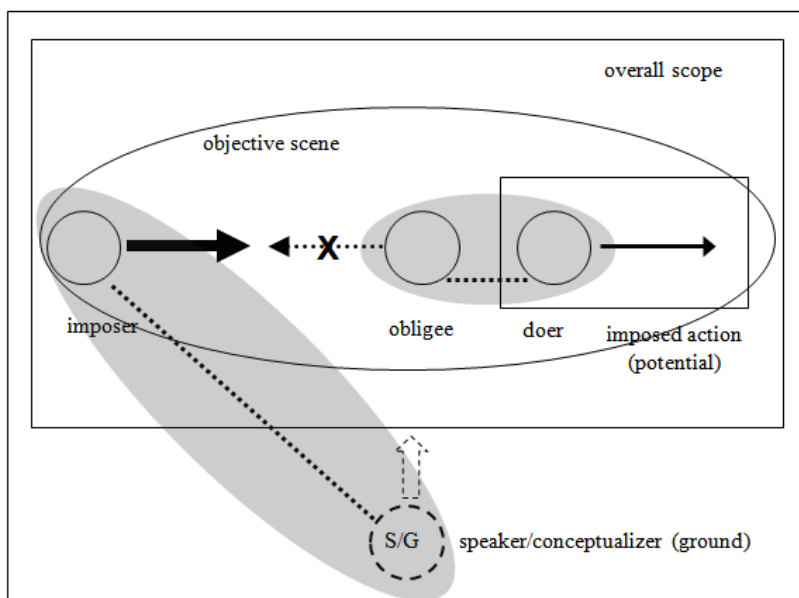

Figure 2: The conceptual structure proposed for English deontic *must* in Pelyvás (2011b).

Scopes and grounding have been added. The arrows of different weights mark the (strengths of) the forces associated with the participants (*'imposer'* and *'obligee'*). The dotted lines mark correspondence: the same participant in alternative roles, our chief concern here.

The deontic scene, illustrated in Figure 2, has two hidden correspondences:

● The imposer of the obligation is normally the speaker;

● The obligee (the subject of the sentence containing the modal) appears in two different roles. One is an agent-like role of performing imposed potential action, the other is the one who receives the order to do something. But, unlike in the standard 'billiard ball model' of a transitive clause (cf. Langacker 1999: 24), where the participant in the middle is entirely passive, the *obligee* has an active role as well: exerts a relatively weak counterforce to the strong force associated with the obligation[6].

If we associate the revealed roles of the conceptual structure with syntactic cases in the organization of the clause, the prediction is that either role can be 'grammaticalized': Nominative case would grammaticalize the agent-like (doer)

---

[6] This makes sure, among other things, that the action remains potential rather than actual—a factor that Sweetser's analysis cannot explain.

role, and a Dative would mark the somewhat active but subordinate role of the 'obligee', with his/her reluctance to perform the action. In English we only have the Nominative, but in a number of languages, including Hungarian, there is (also) a Dative subject option, as seen in (13)[7].

(13) a. *Neked* ki   kell tisztítani/(od)   a   cipődet.
           You-dat (out) must clean-inf-*(2nd.sing)*   the   shoes-acc.-2nd. sing.poss.acc

       b. *Te*   ki   kell, hogy tisztítsd   a cipődet.
           You-nom   (out) must that clean-2nd.sing.subj the shoes-2nd.sing.poss.acc
       'You must clean your shoes.'

The correspondence and the differences in the roles that I have described here are largely hidden in the organization of the clause, although they may not be without significant consequences in the organization of larger units, such as ICM's, scenes or scenarios.

3. Construal: Idealized Cognitive Models in communication

3.1. Conflicting models

Since the creation of an ICM, as we have seen, is largely a matter of construal, there is nothing surprising about situations when the participants of a conversation construe the 'same' situation in terms of different models. This can often lead to misunderstanding, conflicts or frustration in the conversation, especially if the models turn out to be incompatible with each other. In literary texts the author can use such situations as a source of (often black) humour. (The following discussion is based on Pelyvás (2001), in *Officina Textologica* 5.) Consider the following passage from Joseph Heller's *Catch 22*:

(14) ("I'm not joking," Clevinger persisted.)
      "They're trying to kill me," Yossarian told him calmly.
      "No one's trying to kill you," Clevinger cried.
      "Then why are they shooting at me?" Yossarian asked.
   5 "They're shooting at everyone," Clevinger answered. "They're trying to kill everyone."
      "And what difference does that make?"
      ...

---

[7] Some languages may even be more sensitive to these conceptual differences. Romanian, for instance, would only have a (conversational) Dative variant if the subject is +HUMAN, i.e. capable of exerting such a counterforce. For a discussion of how relationships change in the epistemic domain, cf. Pelyvás (2011b).

> "Who's they?" [Clevinger] wanted to know. "Who, specifically, do you think is trying to murder you?"
>
> "Every one of them," Yossarian told him.
>
> 10 "Every one of whom?"
>
> "Every one of whom do you think?"
>
> "I haven't any idea."
>
> "Then how do you know they aren't?"
>
> "Because . . ." Clevinger sputtered, and turned speechless with frustration.
>
> 15 Clevinger really thought he was right, but Yossarian had proof, because strangers he didn't know shot at him with cannons every time he flew up into the air to drop bombs on them, and it wasn't funny at all. ...

Clearly, there is nothing humorous about the situation: this is war, the characters are under the constant pressure of being in danger of violent death. Additional tension is provided by the repeated increase of the number of missions they have to fly, to mention just the most important ingredients (lines 1, 6, 15, 17). Yet Heller somehow manages to turn this situation funny: not for the participants, who are on the verge of a breakdown, but for the reader set apart from them and observing these developments.

To find the source of humour in this unfunny situation, we have to return to the participants: to the tension palpable between them. They both have a strong urge to communicate something really important to them but for some reason they cannot come to terms with each other. They feel this and are frustrated (lines 1, 6, 14) but cannot understand or untangle the situation. This is reflected in the obvious contradiction between lines 3 and 5, in the snappy exchanges, in Clevinger's frustration and Yossarian's puzzlement in the last lines.

The key to the situation is that the two participants experience and conceptualize the situation in two entirely different cognitive models. It is extremely difficult if not impossible to 'step out' of a cognitive model, since 'there is nowhere else to go'. The further the models are apart, the more difficult understanding will be.

The excerpt reveals that Yossarian thinks in terms of the ICM of *murder,* while Clevinger uses the model of *war*. Despite the apparent similarity, the distance between the two models are great, even though Yossarian posits the naive question in line 6: *'And what difference does it make?'*

In the ICM of *murder* there prototypically is a personal element. Murderer and potential victim often know each other well and the murderer has some personal motive of anger, jealousy, hate or potential gain. The words *they, me, try, kill* in Yossarian's lines refer to these elements, duly challenged by

Clevinger in lines 8 and 10 relying on the model of *war,* where '*they are trying to kill everyone'*. This is turned inside out by Yossarian in line 14, finally nailing his opponent.

The ICM of *war* is totally impersonal. In modern warfare the troops hardly see each other, soldiers are trained not to think of the enemy as human and terms like *liquidate, annihilate, pacify* or *mop up* are used instead of the traditional terms. This is especially true of aerial warfare, which is frequently referred to as 'clean'. Thinking in terms of this model, Clevinger frequently uses the terms *everyone, no one,* and this is what leads to deeper meaning of the contradiction *'No one's trying to kill you* [because] *they're trying to kill everyone'*.

Can the characters remain sane in a crazy situation? Heller's ingenuity makes sure that they cannot, and lines 15 to 17 clearly betray this. Yossarian appears to entertain both cognitive models simultaneously: when *they shoot at him,* that is *murder*, but when he *flies up into the air to drop bombs on them,* that is because of *war*, a natural thing.

This section is an example of how the choice of the cognitive model applied to make sense of a situation can determine the success of communication at higher levels. My aim was to show that the cognitive principle of construal can affect coherent communication. Conflicting or incompatible models in the minds of the partners (or sometimes of one person) can make communication extremely difficult if not impossible but can be an excellent source of humour for the reader of a literary piece (who is of course not part of the situation).

3.2. An inside view on the creation of ICMs: psychotic narration

Making sense of a situation is harder work than would appear at first sight. After perhaps a brief period of 'tuning in', nearly all language users are capable of working out an ICM which is compatible with what is seen or heard. (Just think of what happens when you sit down to watch a film that has been running for a few minutes.)

Experiments conducted by Chaika and Alexander (1986) prove that such a task can be very difficult if not impossible for psychotic patients. (Our discussion of the *'ice cream stories'* is based on Pelyvás (2003) in *Officina Textologica* 9, for a fuller English version see Pelyvás 1996.)

To determine to what extent psychotic patients were able to create a coherent cognitive model of a simple scene, they made a short (2 minute) video story of how a little girl obtains money from her parents and buys ice cream[8]:

---

[8] The experimenters had to be very careful in designing the story, as it cannot include anything that could potentially upset the patients. Psychotic patients lack the ability of normal subjects to detach themselves from a situation (objective viewing arrangement) that we referred to as an essential ingredient of the humour of the excerpt taken from *Catch 22.*

(15)   The first scene pans a shopping center, with the camera gradually closing in on a little girl looking through the window of an ice cream store.

The next scene shows a woman setting table, with the same girl walking in and asking, 'Mommy, can I have some ice cream?' The mother answers (gently), 'No, honey, it's too close to suppertime.'

Then a man walks into the house. The child goes up to him, they greet each other, then she asks, 'Daddy, can I have some ice cream?' The father looks into the camera with a grin, and his hand moves towards his pocket.

The next scene shows the child entering the ice cream store, leaning against the counter as she waits fidgeting.

(Then she buys a very large double grape ice and leaves the store.)

(pp. 310-311, abridged)

Even a quick glance at the excerpts from psychotic narratives quoted below can convince anyone that some of the psychotic narratives do not tell the story at all cf (16).

(16)   Okay. I was watching a film of a girl and um s bring back memories of things that happened to people around me that affected me during the time when I was living in that area... (psychotic)

Others do, to some extent, but with great deficiencies in the attempt to create a coherent cognitive model. On point of special significance in the story is the part where there is a gap in the video: *the father's hand moves towards his pocket and then the girl returns to the store and buys an ice cream.* The control group had no difficulty in bridging the gap: *the father must have given the girl some money,* but the task proved too difficult for most of the psychotic subjects. A good example for this is (17).

(17)   ... and I noticed a little girl looking into the window and I guess he walked back into the store and then a [kif] thing switched where the girl was at home and I dunno asked her mother for something and she had a kni- got a little memory lapse there. Then it switched again and her father came in...(psychotic)

The more severe cases even had problems identifying objects/participants and the basic relationships among them, the very first step in creating an ICM. This is evident in (18).

(18)   I saw a little girl who was moving a counter for some reason and I don't know what the heck that was about. She was pressing

against it okay. In the beginning I saw a white car with a red vinyl top and then this little girl was looking in the store was looking in the trash can or something and then she turned around and she went on she talked to her mother and her father and neither one was listening to her... (psychotic)

The subject begins by misinterpreting the girl leaning against the counter as an attempt to move or push it, though admitting that (s)he cannot make sense of this relationship. Then the attempt at 'tuning in' is obviously given up when (s)he starts listing details that are discarded as irrelevant at the beginning without difficulty (the car, its vinyl top, etc.) by the healthy control group.

The most interesting detail in this narrative is probably the *trash can*. It is normally taken for granted in the literature that the participants (prototypically 3-dimensional objects) have greater integrity in the ICM than relationships, at least in the sense that they are conceptualized as existing independently of the situation. This narrative suggests that it may not always be so. Even without actually seeing the video we can argue that the psychotic narrator would not have identified the ice cream containers as trash cans if (s)he had understood that the ICM was one of buying ice cream.

Objects may acquire their proper conceptualization from the relationships that they participate in. This appears to be an even more fundamental property of construal than the selection of an appropriate argument structure for verbs.

3.3. The impossible scenario

In the previous section we have seen something of what it takes to create an Idealized Cognitive Model of a situation through the example of psychotic patients, who are often not capable of the mental operations of distancing themselves away from a situation, of finding the proper scope for the narrative, of identifying participants or simple relationships holding among them, or of bridging gaps in the network. These operations come so naturally to the normal speaker that (s)he is even capable of making sense of scenarios that 'do not make sense'.

Even little children can effortlessly understand and enjoy the cartoon scene in which a character, having reached the brink of a precipice, walks on whistling to himself—until he looks down, gets frightened, and has the nasty fall. This is turning the natural course of events round, making believe that the laws of gravity somehow depend on our observation.

Sometimes we encounter impossible scenarios and we can not only 'accept' them but can also understand the hidden meanings that they are meant to convey. Here is an example of one of István Örkény's grotesque *One Minute Stories*. The discussion is based primarily on Pelyvás (2008) in *Officina*

Textologica 14, but see also the other contributions in that volume, especially Kiss S. (2008) and Csűry Andrea (2008).

> (19) István Örkény: The Death of an Actor
> The popular actor Zoltán Zetelaki collapsed and lost consciousness on a street just off Rákóczi Road early this afternoon. Passers-by called an ambulance and rushed him to a nearby clinic. Despite the application of the latest advances known to medical science including the use of an iron lung, all efforts to revive him were in vain. At 6.30 in the evening, after lengthy agonies the celebrated Thespian died and his remains were transferred to the Institute of Anatomy.
> Despite this terrible misfortune tonight's performance of King Lear proceeded as usual. Though a few moments late and **looking rather worse for the wear**—in Act 1 here and there he had to rely on the prompter—Zetelaki **gradually revived and by Act 5 he was so convincing as the dying king** that the audience gave him a standing ovation.
> After the performance Zetelaki was invited out to dinner but he declined. 'Thank you very much,' he said, 'but I've had a rather trying day.'
> (Translation by Judith Sollosy, emphases are mine)

Sudden death of the actor in real life would make the offered scenario impossible. The reasons that this is not so for Örkény are quite complex and create an artistic effect in a complex interaction that cognitive linguistic theory calls *conceptual integration* or *blend* (cf. Coulson and Oakley 2000, Grady et al. 1999). Admitting that we are now approaching the somewhat unfamiliar grounds of literary analysis, the linguist can observe at least the following factors in interaction:

● It is customary in the modern world to constantly spy on the private lives of celebrities and make all detail visible to the public. Arguably the story satirizes on the appropriateness and reliability of such information by juxtaposing the two parts of the story.

● Even the average man often wonders about the background or source of artistic inspiration. It is somehow felt that an actor must have some sort of personal experience of the situations before (s)he can convey them convincingly to an audience. If this is true, death could only be performed well after really experiencing dying, i.e. it would be impossible unless the scenario developed above could be real (or the dogma about real experience is false).

● With Paragraph 3 we can witness the gradual development of a conceptual blend which integrates the elements of *real death and death on the stage, exhaustion as metaphorical death, being not quite up to the mark, and of the daily routine of an actor's work,* culminating at the point when death on the stage blends with death in real life[9]. Ironically, this is the phase most appreciated by the audience. The finishing lines of the story tell us that this impossible scenario needs to be repeated as a routine day after day after day in life.

References

1. The discussion above is based on the following papers by the author in the volumes of *Officina Textologica*:

Pelyvás Péter 2001. Az idealizált kognitív modellek szerepe a mondat- és szövegalkotásban.[The role of Idealized Cognitive Models in sentence structure and text construal] In: Petőfi S. János – Szikszainé Nagy Irma (ed.) *Officina Textologica 5.* Grammatika – Szövegnyelvészet – Szövegtan. Debrecen, 2001. 11-26.

Pelyvás Péter 2002.   Szórend, grammatikai funkciók és téma–réma szerkezet a kognitív grammatikában (Kognitív alapú angol–magyar összevetés). [Word order, grammatical functions and theme-rheme structure in cognitive grammar (A comparative study of the English and Hungarian languages following a cognitive approach.)] In: Petőfi S. János–Szikszainé Nagy Irma (ed.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 19–34.

Pelyvás Péter 2003.   A tematikus progresszió és annak hiánya angol nyelvű pszichotikus szövegekben. [Thematic progression and its absence in English psychotic narrations.] In: Petőfi S. János–Szikszainé Nagy Irma (ed.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 17–32.

Pelyvás Péter 2005. Rejtett megfelelések (korreferencia-viszonyok) az angol módbeli segédigék konceptuális szerkezeteiben. [Hidden correspondences (coreference relations) in the conceptual structures of the English modals]. In: Petőfi, S. János – Szikszainé Nagy Irma (ed.) *Officina Textologica 12.* A korreferencialitás poliglott megközelítése Debrecen, 2005. 73-81.

---

[9] The translation very appropriately uses *revive* twice, both in the literal and the metaphorical sense. In the Hungarian original the two expressions are *életre kelt* and *magára talált*.

Pelyvás Péter 2008. Forgatókönyv-építési stratégiák. [Strategies for building scenarios.] In: Dobi Edit (ed.) *Officina Textologica 14.* A forgatókönyv mint dinamikus szövegszervező erő. Debrecen, 2008. 71-78.

Other references

Chaika, Elaine and Alexander, Paul 1986. The Ice Cream Stories: A Study in Normal and Psychotic Narrations. *Discourse Processes* Vol. 9, No.3, 305-28.
Chomsky, Noam 1964. Current issues in linguistic theory. In: Fodor, J.A. and Katz, J.J. (eds.) *The Structure of Language: Readings in the Philosophy of Language.* pp. 50-118. Prentice-Hall.
Coulson, Seana and Oakley, Todd 2000. Blending basics. *Cognitive Linguistics* 11-3/4. 175-196.
Fillmore, Charles 1968. The Case for Case In: Bach, E. and Harms, R.T. (eds.) *Universals in Linguistic Theory.* New York: Holt, Rinehart & Winston.
Fillmore, Charles 1977. The Case for Case Reopened. In: Cole, P. and Sadock, J. M. (eds.) *Syntax and Semantics 8.* New York: Academic Press.
Grady, Joseph E., Oakley, Todd and Coulson, Seana 1999. Blending and metaphor. In: Gibbs, R.W. and Steen, G.J. (eds.) *Metaphor in Cognitive Linguistics.* Amsterdam: Benjamins, 101-124.
Langacker, Ronald W. 1987, 1991a. *Foundations of Cognitive Grammar I, II.* Stanford, CA: Stanford University Press.
Langacker, Ronald W. 1999. *Grammar and Conceptualization.* Berlin and New York: Mouton de Gruyter.
Newmeyer, Frederick J. 1991. Functional explanation in linguistics and the origins of language. *Language and Communication* Vol. 11, No. 1/2 3-28.
Pelyvás Péter 1996. *Subjectivity in English.* Frankfurt am Main: Peter Lang.
Pelyvás Péter 2011a. Motivation in English *must* and Hungarian *kell*. In: Panther, Klaus-Uwe and Günter Radden (eds.) *Motivation in Grammar and the Lexicon.* Amsterdam and Philadelphia: Benjamins: 171-190.
Pelyvás, Péter 2011b. Epistemic Modality: a Choice Between Alternative Cognitive Models. In: Pârlog, H. (ed.) *B.A.S. (British and American Studies) 2011.* Timişoara: Hestia. 247-259.
Sweetser, Eve. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge: Cambridge University Press.
Tolcsvai Nagy Gábor 2001. *A magyar nyelv szövegtana.* Budapest: Nemzeti Tankönyvkiadó.

Literary sources

Heller, Joseph: Catch-22. Corgi Books, 1961.

Örkény István: A színész halála. In: *Egyperces novellák,* Magvető, 1977. [The Death of an Actor, translated by Judith Sollosy].

3.
Co-reference[1]
ANDREA NAGY—FRANCISKA SKUTTA

1. Co-reference means the relation ensuring the textual connectivity of two or more linguistic elements with different verbal manifestations which, according to the interpreter, refer to the same entity of the given text world. As to the analysis of the relation between the elements which are considered to be coreferential, i.e. co-reference analysis, JÁNOS S. PETŐFI and EDIT DOBI (1998: 238) write the following:

> **"The *analysis of co-reference* must take into consideration the relation of two levels: (a) one is the interpreter's mental image of the 'world fragment' represented in the analysed text, with its entities (i.e., with the mental images of persons, objects, concepts, attributes, events, etc.); (b) the other is the physical manifestation of the analysed text, with the text components interpreted syntactically / semantically by the interpreter. A part of these text components refers to entities of the supposed 'world fragment'; the other part expresses diverse statements about the entities referred to.**
>
> **A *co-reference relation* ('together-referring relation') is nothing else than a sort of 'affinity' relation between the text components supposed to refer to the same entity of the presumed world fragment. This relation can be 'identity' (such a relation exists between — either identical or different — text components which presumably refer to the same entity of the presumed world fragment); it can be in a certain sense that of 'possessor-possession' (this holds, among others, for two text components, one of which refers to a person or an object, while the other refers to an entity which can be considered as one belonging to this person or object), but it can be of several other types as well."**

2. A particularly important question in the case of co-reference analysis from a text linguistic / textological point of view is the most effective way of representing explicitly the coreferential elements and their mutual relations in a given text, as well as the deducible information based on linguistic and general knowledge about the world, both of which are necessary for revealing the coreferential relations in the analysed text.

"A polyglot research program in textology / text linguistics" has provided for these question-raising theoretical and practical problems a possible solution, namely the co-reference analysis by JÁNOS S. PETŐFI (1997: 24-38), (1998: 15-31), a highly inspiring model for co-reference analyses of texts belonging to different types.

In fact, this basic method of analysis has made it possible to compare directly the results of the co-reference analysis of texts belonging to different types of texts.

In our essay, we shall first present the model of JÁNOS S. PETŐFI (1998: 15-31); our presentation is based on his study published in *Officina Textologica 2* and will be followed by that of related articles on co-reference.

The sample text analysed in the orientating article is a biblical text extracted from *New Testament*:

> **Ve:** Mt. 9, 9-13. The Calling of Matthew[2]
>
> [9]As Jesus passed on from there, he saw a man called Matthew sitting at the tax booth. He said to him, "Follow me." And he rose and followed him.
>
> [10]And as Jesus reclined at table in the house, behold, many tax collectors and sinners came and were reclining with Jesus and his disciples. [11]And when the Pharisees saw this, they said to his disciples, "Why does your teacher eat with tax collectors and sinners?" [12]But when he heard it, he said, "Those who are well have no need of a physician, but those who are sick." [13]Go and learn what this means, "I desire mercy, and not sacrifice." For I came not to call the righteous, but sinners.

János S. Petőfi resumes the expedient steps of co-reference analysis as follows:

"*First*, I assign — giving up the biblical verse numbering — codes in square brackets (see for example **[C01]**) to the text sentences to identify them. The letter 'C' of these codes stands for the expression 'first-degree macro-**c**ompositional unit', while the number indicates that given to a text sentence in the order of sentences in the text. A text sentence is defined as a unit beginning with a capital letter and closed with a full stop, a colon, an exclamation mark or a question mark.

As the *second* step, I complete each text sentence with verbal expressions (in italics) which refer to elements / relations deducible from the sentences themselves, from their verbal context and / or from our general knowledge about

---

[2] English Standard Version, http://www.biblegateway.com. – The original analysis uses the Hungarian translation of the *Bible*.

the world in a way that each text sentence shall constitute an independent 'information unit', while coreferential relations inside and between the text sentences must be expressed explicitly. In order to indicate that in the analysed text sentence I consider an expression consisting of more than one elements as a single component, I use the concatenation symbol "^". I assign to the text sentences completed with verbal expressions a code completed with the symbol "/&vb" (see for example [C01/&vb]).

As the *third* step — under the title "Comment"— I analyse both the inner syntactic structure and the coreferential relations of each text sentence completed with verbal elements. (However, for the sake of simplicity, I do not deal with articles, negative particles, time adverbials, tenses, and conjunctions.) […]

As the *fourth* step, I determine which elements of the text sentence completed with verbal expressions should be represented by co-reference indices, I create these indices adding a comment to them; for indexation, I use a code consisting of the letter "i" and of a two-digit number (see e.g. **i03** or **i11**).

As the *fifth* (and last) step, I finally create for each text sentence its variant completed with co-reference indices. After their original code, I assign to the text sentences completed with co-reference indices the symbol "/&ind" (see e. g. [C01/&ind]). I create the variants with co-reference indices based on the analysis of the text sentences completed with verbal expressions in the following way:

− I place co-reference indices each time in square brackets after the expression / word / suffix to be marked by them,

− I put co-reference indices in bold type when they represent nominal expressions at their first occurrence and I put an equal mark before them,

− I use normal characters for co-reference indices in all other cases, without an equal mark." (PETŐFI 1998: 15-17)

3. We shall now illustrate the analysis described above. First, we determine the text sentences of the sample text and assign a code to each of them.
Ve[C00]–[C12]:

[C00]*Mt. 9, 9-13. The Calling of Matthew*

[C01]As Jesus passed on from there, he saw a man called Matthew sitting at the tax booth. [C02]He said to him, [C03]"Follow me". [C04]And he rose and followed him. [C05]And as Jesus reclined at table in the house, behold, many tax collectors and sinners came and were reclining with Jesus and his disciples. [C06]And when the Pharisees saw this, they said to his disciples, [C07]"Why does your teacher eat with tax collectors and sinners?" [C08]But when he heard it, he said, [C09]"Those who are well have no need of a physician, but those who are sick. [C10]Go and learn what this means, [C11]'I desire mercy, and not sacrifice.' [C12]For I came not to call the righteous, but sinners.

As the next step, we have to complete the text sentences with the information deducible from the verbal context and from our general knowledge about the world. Let us take text sentences **[C01]–[C03]** as an example.

[C01] As Jesus passed on from there, he saw a man called Matthew sitting at the tax booth.

[C01/&vb]:

As Jesus passed on from there *(from the place where Jesus previously was)*, he *(Jesus)* saw a^man^called^Matthew *(who was)* sitting at the^tax^booth.

[C02] He said to him,

[C02/&vb]:

He *(Jesus)* said *(the following)* to him *(to the man called Matthew)*,

[C03] "Follow me".

[C03/&vb]:

(You, the man called Matthew) Follow me *(Jesus)*.

After the detailed analysis under the title "Comment", the following co-reference indices can be assigned to the elements of the text sentences made explicit by verbal expressions:

i01 = Jesus,

(**i02** = *the place where Jesus previously was*),

i03 = the man called Matthew,

i04 = the tax booth.

As the final step of the analysis, we assign the above co-reference indices to the expressions of the text sentences, thus replacing the explicit verbal information.

Ve/&ind: *Mt. 9, 9-13. The Calling of Matthew* [C01]–[C03]

[C01/&ind]:

As Jesus**[=i01]** passed[i01] on from there*(i02)*, he[i01] saw[i01] a^man^called^Matthew**[=i03]** sitting[i03] at the^tax^booth**[=i04]**.

[C02/&ind]:

He[i01] said[i01]**[C03]** to him[i03],

[C03/&ind]:

Follow[i03] me[i01].

The analytical steps continued in this way, i.e., by examining text sentences one by one, result in the complete list of co-reference indices in a given text. The following list shows the co-reference indices of the text analysed here:

i01 = Jesus,

(**i02** = *the place where Jesus previously was*),

i03 = the man called Matthew,

i04 = the tax booth (the place at the tax booth),

i05 = the house of the man called Matthew = the house(i03)[=i05],

i06 = the many tax collectors (coming to the house of Matthew),

°**i07** = the many sinners (coming to the house of Matthew),

***i08** = the disciples(i01)[=i08] of Jesus (in the house of Matthew),

i09 = table/tables (in the house of Matthew),

i10 = the [’(*the many tax collectors and the (many) sinners*) with Jesus and his *(Jesus’)* disciples together (*in the house of the man called Matthew*) reclining at table)’ event-seeing] Pharisees,

i11 = those who are well (in general),

i12 = those who are sick (in general),

i13 = a physician (in general),

i14 = (unspecified) mercy,

i15 = (unspecified) sacrifice,

i16 = the (unspecified) righteous.

The distribution of coreferential elements in the text sentences can also be represented in the form of a table clearly indicating the co-reference relations in the given text. The co-reference relations of text sentences **[C01]–[C03]** analysed above are represented by the following table:

	[C01]	**[C02]**	**[C03]**
i01	**Jesus,** passed on, he, saw	he, said	me
i02	*(from the place)*		
i03	**the man called Matthew,** sitting	to him	Follow
i04	at the **tax booth**		

♦

It is then along the main lines of this orientating article that further investigations in co-reference have been carried out by the contributors of "A polyglot research program in textology / text linguistics". Relevant studies are to be found in several issues of *Officina Textologica*, the most important volume in this respect being precisely the one — *Volume 2* (1998) — containing Petőfi's leading article and eleven other studies on aspects of co-reference. In fact, under the title *Coreferential elements, co-reference relations*, the volume offers analyses of a wide range of phenomena concerning the realization of co-reference in different text types represented either by complete short texts or by extracts from longer ones. This first thematic collection of articles was then followed by a so-called "discussion" volume —*Volume 4* (2000) —, which

raises some more theoretical issues related to co-reference. While these two volumes analyse coreferential elements and relations exclusively in Hungarian texts (or, occasionally, in texts translated into Hungarian), the six essays of *Volume 12* (2005) examine problems of co-reference in a contrastive approach, comparing Hungarian with English, French and German, respectively. Finally, several other issues of *Officina Textologica*, though not necessarily focussing on co-reference, contain articles which occasionally touch upon questions of connectivity and cohesion. Since it would be impossible to give a detailed presentation of each study in the framework of this short essay, and in order to avoid repetition, we shall discuss certain *problems* of co-reference analysis rather than present individual articles. Owing to the great variety of coreferential phenomena and of analytic methods, we shall concentrate on *Volume 2* and examine how different text types may influence, or even command, the linguistic formulation of co-reference relations.

As we know, texts can be classified into types, which are best defined as a set of features selected according to different criteria, both linguistic and pragmatic. In fact, from a typological point of view, the texts described in the eleven articles that follow the Petőfi model in *Volume 2* can be divided in the first step into non-literary and literary texts, each of these two general types being represented here by specific genres.

Non-literary texts — conceived for usage in various every-day situations and interactions — are exemplified in this volume by certain genres which are embodied either in texts with usually reduced dimensions, or on the contrary, in texts of considerable length like scientific or scholarly books. To illustrate the former group, the authors quote and analyse such complete short texts as an *encyclopaedia entry* (3, DOBI), a *recipe* (4, DOMONKOSI), a *business letter* for advertising purposes (5, CS. JÓNÁS), and a *newspaper article* (11, SZIKSZAI NAGY), whereas the latter type appears in one article on a *scholarly text*, represented by an extract from a book on literary theory (6, SKUTTA). A special case of non-literary texts, namely a new type of *multimedia product*, the CD-ROM with its *Help* feature, is also dealt with (8, BENKES—VASS).[3]

Literary texts proper appear in three articles, one analysing a *poem* (10, B. FEJES), another a short extract from a *novel* (9, BÉKÉSI), and the third an "entry" from a so-called "*dictionary-novel*" (12, TOLCSVAI NAGY). Finally, two borderline cases of literature are treated in two studies respectively: a *folk-tale* (7, TUBA) and a biblical text, an extract from *New Testament* (2, BODA—PORKOLÁB).

---

[3] The number in brackets (preceding here the name of the author and later used without a name) indicates the number in the order of articles in *Volume 2*. See the list of articles below.

As it may be seen from the above list, the corpus used for the analysis of co-reference is sufficiently varied for possibly representing a considerable amount of phenomena which can be examined from the point of view of textology / text linguistics. It seems that a comprehensive survey of those phenomena — as the one by PETŐFI—DOBI at the end of *Officina Textologica 2* (pp. 238-261) — has to answer the following questions:

(1) What kind of *world entities* can be referred to by means of specific referential units of the text?

(2) What sort of elements can be used as *referential text units*?

(3) What kind of *co-reference relations* can be identified, in other words, what sort of referential differences may be included in the notion of co-reference?

(4) Which is the most effective way of *representing* coreferential elements and co-reference relations?

In what follows, we shall discuss these four questions one by one, with the answers suggested by the presentation of PETŐFI—DOBI and the analyses in *Volume 2*. It is important to bear in mind, however, that an exhaustive description of all relevant phenomena would need an investigation on a much larger corpus.

(1) *World entities* referred to by linguistic means are either countable or uncountable and both classes contain further subdivisions.

Countable world entities are persons, objects, etc.:

(a) numerically well-defined and clearly identifiable:
Jesus (1), Emperor Nero (3), Thomas Mann (10), Levin (9), Vronsky (9)
the king (7), the party leader (11), the mother (9), the coachmen (3)
the chicken (4), the cock (7)
the house of the man called Matthew (1), the red VW Golf (5), the gold coin (7)
the large and high mountain where the angel led John (2), the Holy City (2)
the Roman Empire (3), the *Circus Maximus* (3)
the Renaissance (6)
the two aspects (6), the three main parts (8)

(b) numerically well-defined but not clearly identifiable:
a royal family (7), the 150 to 200 thousand people watching the race (3)
an unspecified CD-ROM (8)

(c) numerically indefinite but clearly identifiable:
the tax collectors and the sinners in the house of Matthew (1), the servants (7)
tables in the house of Matthew (1), the products of our firm (5)
the written records on chariots races (3), the works of the classical authors (6)

(d) numerically indefinite and not clearly identifiable:
people (7), those who are well (1), those who are sick (1)
the loveliest fruit trees (5), the chariots (3)

unspecified words (8), the conditions of the genesis of (literary) works (6)

Uncountable entities are:

(a) of a definite quantity and clearly identifiable:

the length / width of the *Circus Maximus* (3)

(b) of a definite quantity but not clearly identifiable:

150 g of mushrooms (4), 1 spoonful of parsley (4)

(c) of an indefinite quantity but clearly identifiable:

the material of the walls of the Holy City (2), the water in the well (7)

(d) of an indefinite quantity and not clearly identifiable:

glass (2), a pinch of salt (4), water (4)

Entities *different* from the above are of six main types:

(a) objects whose consistency is continually changing:

1 chicken (4), 150 g of rice (4)

(b) abstract notions:

mercy (1), sacrifice (1), literature (6), the idea of duality (6)

(c) linguistic / metalinguistic entities (titles, names, forms of address):

the encyclopaedia entry "Chariots race" (3), Dear Mrs. Szegedi (5), Princess (7)

(d) qualities:

of childhood (12)

(e) states of affairs:

(sb) asks (sb to do sth) (10)

(f) time, period:

again (11), the time of the tale (7)

(2) Types of *referential text units*

(a) proper names:

Jesus (1), Hans Castorp (10), the Roman Empire (3)

(b) nouns, noun phrases:

the bride (2), the little coin (7), the history of literary research (6), in the dark (10)

(c) pronouns (personal, demonstrative, relative, etc.):

he (12), me (1), with us (10), these (8), in that (8), who (12), which (2)

(d) [in Hungarian] personal possessive suffixes (referring to the possessor):

<u>my</u> name (12), <u>our</u> catalogue (5), the beginning <u>of [it =] the race</u> (3)

(e) adverbs:

thus (2), then (7), again (11)

(f) [in Hungarian] verbal prefixes:

you give it <u>back</u> (7), it flew <u>up</u> (7)

(g) [in Hungarian] verbal suffixes (referring to the grammatical person):

I know (12), he saw (1), you can choose (5), take a seat (10)

(h) infinitive:

to resemble (9), to remember (11)

(i) finite forms of verbs as autonomous states of affairs:

(sb) asks (sb to do sth) (10), (sb never) lies

(j) non-linguistic symbols:

my ~ [for *ideal*, an entry in the so-called dictionary-novel] (12)

(3) Types and limits of *co-reference relations*

(a) complete referential identity (repetition, synonymy, pronominal substitution, indication of the person of the possessor, conjugation, appositive — occasionally defining — constructions):

Jesus, he, me, him, his disciples, your teacher, I (1)

(b) hyponym — hyperonym relations:

hyperonym: all sorts of precious stones — hyponyms: jasper, emerald, topaz (2)

(c) associative relations:

his photo — light (11)

(d) set — subset relations:

the 150 to 200 thousand people watching the race — those making a bet (3)

(e) change of state (of consistency) of the entity referred to:

the chicken, the chicken cut to pieces, stew it, spice it, serve it (4)

(f) metalinguistic correspondence:

chariots race [encyclopaedia entry] — speed contest (3)

(g) reference to text segments:

tax collectors and sinners […] reclining with Jesus and his disciples — this (1)

(h) co-reference between a noun phrase and its central semantic constituent:

the literary work(s) — literature (6)

(i) co-reference ensured by the given communicative channel:

tell us a tale — your words — you speak — we listen (10)

(4) Ways of *representing* coreferential elements and co-reference relations

At the present state of research, it is impossible to give a definitive answer to the question concerning the most effective way of representing coreferential elements and co-reference relations. While based on the representational system used in the orientating article, the analyses of different text types suggest various additional solutions specific to those types of texts. Since the use of indices for representing simple noun phrases (e.g. proper names or a noun preceded by a determinant and interpreted in its concrete meaning) does not seem to cause any difficulty, we shall focus below on certain problematic cases of indexation. In fact, some of those phenomena are specific to the Hungarian language, such as the existence of two conjugations — "subjective" and "objective" —, as well as of verbal prefixes and of personal possessive suffixes (added to nouns), while a considerable number of phenomena are clearly translinguistic, among others forms of address, or reference made either to a complex noun phrase or to one of its constituents only. In what follows, we shall briefly comment on a series of "problematic issues", raised by the analyses in *Officina Textologica 2*.

I. *Types of referential text units*

– *Personal suffixes in the objective conjugation*

These suffixes ensure a double co-reference, i.e., with the subject and with the definite direct object of the verb. Both co-reference indices — the first representing the subject, the second the object, and being separated by the sign — should be marked immediately after the conjugated form of the verb:

The cock [i01] gave [i01 i02] the coin [i02] [to the Princess] (7).

– *Verbal prefixes*

Since Hungarian verbal prefixes can eventually refer to the participants of the action designated by the verb, in such cases they should be supplied with co-reference indices of their own. The following English translation cannot reproduce exactly the original verbal prefixes, which are then represented by the preposition *to* and the adverbial particle *back*:

I [i01]'ll give [i01] it [i02] to you [i05], Princess [i05], but only if you [i05] give [i05] it [i02] back [i01] (7).

– *Personal possessive suffixes*

In Hungarian, nouns can take so-called "personal possessive suffixes", which refer to the possessor and, at the same time, indicate the grammatical person of the possessor referred to in a particular communicative situation. Thus, functionally, they correspond to English possessive determinants (*my*, *your*, *his*, etc.), but since in Hungarian the possessive reference is incorporated in the noun, the latter will be followed by a complex co-reference index, the first member of which represents the referent designated by the noun, and the second refers to the possessor expressed by the suffix:

the king [i04]'s sister [i05 +i04], an elderly princess [i05] (7).

– *Title, name, social rank, forms of address*

Such phrases can be used either to refer to world entities or to quote text segments, but when a co-reference chain occurs in the text, the *title / name / social rank / forms of address* phrases normally refer to world entities:

chariots race [entry] — speed contest [world entity] (3)

She [world entity] was called *Princess* [name], because her [world entity] brother [world entity] was a *king* [rank] (7).

The Princess saw it [the cock = world entity] and said: 'My dear *cock*' [form of address] (7).

Such phrases, when used as "quotations", should be marked with special additional symbols (˚, *, ').

– *Concrete and abstract meanings of a word*

A sort of "type / token" relation holds for words with both concrete (token) and abstract (type) meanings, such as:

a [particular] gold *coin* found by the cock — *coin* ['money' in general] (7).

Such phrases, when used in their abstract meaning, should be marked with special additional symbols (*).

– *Reference made to a text sentence or to a larger text segment*

This type of reference is often realized by means of a demonstrative pronoun (*this*, *that*); indexation can resume the whole sentence / segment referred to by use of a code standing for the macro-compositional unit in question:

[C05]And as Jesus reclined at table in the house, behold, many tax collectors and sinners came and were reclining with Jesus and his disciples.

[C06]And when the Pharisees saw *this***[C05]**, they said**[C07]** to his disciples,

[C07] "Why does your teacher eat with tax collectors and sinners?" (1).

II. *Types and limits of co-reference relations*

– *Change of state of the entity referred to*

Each phase of a change of state can be represented in a complex way, by means of the constant index standing for the entity and the codes (in subscript) of text sentences describing the actual states of the entity:

[C02] *the chicken* [i02] cut to pieces — *the* roast *chicken* [i02$_{C02}$] (4).

– *Set — subset relations*

In order to show that the validity of a predicative statement does not cover the whole set, but extends only to a subset of the whole, an additional symbol may be introduced ("⊂" = 'subset'):

the 150 to 200 thousand people [i07] watching the race — *those* [⊂i07] making a bet (3).

– *Associative relations*

It is possible to conceive different representations of associative relations — as those between *photo* and *light* (11)[4] —, but it is advisable to use one and the same index for the common semantic feature, accompanied by additional symbols, e.g. subscripts.

– *A noun phrase and its central semantic constituent*

In order to show the presence of the same semantic constituent in two formally different noun phrases, it is possible to use the same numerical index accompanied by two different letters:

the literary [j01]^work(s) [i01] — literature [j01] (6).

– *Co-reference indices whose interpretation implies other referential elements*

In such cases an additional symbol (e.g. *) can indicate the presence of another — autonomous — referential element in the text, which is necessary for the interpretation of the referential unit in question:

---

[4] In Hungarian the word for *photo* is a compound containing the word for *light* followed by the one for *picture*: "light-picture". The associative relation is thus quite explicit. – Study (11) uses an indexation different from that of the orientating article, which would be too long to explain in a short comment.

the *Circus Maximus* [i19] — (length [i19][i20*], width [i19][i21*]) (3).
III. *Simple referential* (*not coreferential*) *index*

It is advisable to use a special symbol (e.g. **) for indicating the single occurrence of a referential element which does not enter a co-reference chain:

antiquity [i04**] (3).

♦

Having examined various problems in connection with (1) world entities referred to, (2) referential text units, (3) types and limits of co-reference relations, and (4) the most effective ways of representing coreferential elements and co-reference relations, we shall now turn briefly to the twelve articles of *Officina Textologica 2*, in order to treat some special questions raised by the types of texts analysed in this volume.

By giving an explicit co-reference analysis of a short extract from *New Testament*, the orientating article (1) by JÁNOS S. PETŐFI offers a model applicable for any text type, with minor changes in the representation of co-referential elements and co-reference relations. That is certainly due, among others, to the nature of the analysed text itself: in fact, *The Calling of Matthew*, *Mt. 9, 9-13* is a sort of "unmarked" text, an almost everyday *narrative* sequence with a relatively simple, chronological time structure and with a small number of actors — Jesus, Matthew, the disciples, the Pharisees, the tax collectors and the sinners coming to Matthew's house —, whose multiple textual occurrences (in the form of noun phrases and pronouns) create well-defined co-reference chains. Thus the explicit representation of those chains by means of coreferential indices remains unequivocal throughout the passage (for details, see above).

The authors of Study (2), KÁROLY I. BODA and JUDIT PORKOLÁB, examine another extract from *New Testament*, namely a passage presenting the New Jerusalem in the *Revelation to John* (*Rev. 21, 9-23*). Though set in a narrative framework — one of the seven angels coming to John, carrying him away to a high mountain and showing him the Holy City — John's vision is rendered in a *descriptive* sequence, with a spatial organization of the City and its parts, the wall, the twelve gates and the street. Whereas in the *Calling of Matthew* co-reference chains represent first of all human beings performing actions in time, indices identifying the elements of John's vision in the *Revelation* stand for static objects characterized by measures, forms, materials and colours. Since the vivid description of the radiant city implies emphatic repetition of the names of its parts, the extract lends itself to an analysis slightly different from that of the orientating article: the authors have invented a frequency matrix for representing the textual occurrences of referential units.

Study (3) by EDIT DOBI deals with a special type of short text, namely an *encyclopaedia entry*, 'chariots race', extracted from the *Hungarian Larousse Encyclopaedia*. This entry combines characteristics of narrative and descriptive texts, for it *describes* a social *event*: in fact, the chariots race, though occurring in time, will not turn into a *story*, and thus the presentation has to focus on its general features and circumstances. Given the concise *informative* and *definitive* nature of encyclopaedia entries, the analysis of co-reference relations is concerned with several specific problems, among others the relationship between the referential and the metalinguistic interpretations of a noun phrase and, in the latter case, the use of a special symbol; divisions of the entry and definitions; abbreviations and non-linguistic symbols. Questions of a more general validity concern the identification and the separation of text sentences, especially of those that form the typical introduction to the entry. Following the definitive nature of this type of text, the concept of "partial co-reference" and the possibility of its representation are also examined.

In some respects, Study (4) by ÁGNES DOMONKOSI raises issues similar to those relevant for the textual organization and the visual manifestation of an encyclopaedia entry. Here the selected text, a *recipe*, 'chicken with rice' (from Lukács Túrós: *Cookery Book for Maidens and Women*), is also a short text (of about the same length as the entry) with a title and a typographical distinction of text segments, so that these aspects of the text must be taken into consideration in the course of referential indexation (e.g. metalinguistic interpretation, the coreferential status of 'ingredients'). Furthermore, the recipe also describes a process occurring in time, by giving instructions for a series of operations to be performed for the realization of a goal. However, as opposed to the entry — and in fact, to all other texts examined in this volume —, the recipe refers to world entities continuously changing in the process of cooking, which raises the problem of 'identity' / 'difference' and, consequently, of coreferential indexation. The solution suggested here marks each phase of the alteration of the entity prior to its state under examination.

With Study (5) by ERZSÉBET CS. JÓNÁS, we turn to a type of text rather different from the previous types in one respect, namely as far as reference to the text recipient is concerned. In fact, the first three texts analysed in the volume make no explicit reference to the reader. As for the recipe, its potential readers / performers are at least implied in the text by means of verb forms (1st person plural). On the contrary, the text analysed here, a *business letter* written for advertising purposes (i.e., each copy of it) is clearly addressed to a particular recipient, using names, typical formulas of politeness and often the imperative of the verbs, with which the sender tries to elicit a positive response to the offer. The analysis therefore takes into account the sociological and communicative aspects of the business letter, as well as the convincing power of the text. The

description of co-reference relations faces some text specific problems, such as the different representations of and the constant feedback to the addressee.

As compared to the previously analysed texts, in which, for the most part, new arguments introduce new world entities, the text examined in Study (6) by FRANCISKA SKUTTA represents a type which aims at describing, in *specialist language*, different aspects of one and the same object, often an abstract domain. The author analyses a short extract — the first two paragraphs — from a comprehensive book on literary theory (Elemér Hankiss: *Literary Work as Complex Model*), and wishes to show some genre-specific characteristics of scholarly writing, which have an impact on co-reference relations. It occurs that in *argumentative* texts referential units forming a co-reference chain are often linked by partial co-reference of different types (set / subset, general / particular, abstract / concrete). Furthermore, the text abounds in complex noun phrases of which only one element enters a co-reference chain, a problem that needs special attention: it is namely advisable to elaborate a way of representing the central and peripheral elements by two distinct, though related, indices.

Study (7) by MÁRTA TUBA analyses an extract from a narrative text rather different in style from the previously quoted texts: it is namely a Palóc *folk-tale* (from the North of Hungary), *The Cock's Gold Coin*, told in a somewhat crude popular language. The syntactic structures of text sentences being relatively simple, with a certain number of repetitions characteristic of folk-tales, the author focuses on some other text-specific features. Since the strongest organizing factor of this type of narrative is undoubtedly the flow of time, the chronological succession of events, all adverbials of time are represented by one common index: "i00", which is a new element added to the analysis suggested in the orientating article, showing the key role of time in the tale. The representation also gives an account of the frequent double functioning of the same expression (usually a noun phrase), used either as a form of address, or as the indication of a social rank — in both cases, the same index appears with a supplementary symbol to distinguish between the two interpretations.

Study (8) is an exception among the articles of this volume in that it deals with a type of communication that is not (exclusively) linguistic. The authors, RÉKA BENKES and LÁSZLÓ VASS, examine, within the framework of semiotic textology, a multimedia product, *Hungarian Bookshelf* CD-ROM, and in particular, its *Help* feature. This medium is composed of verbal and non-verbal — iconic — elements, the totality of which can be considered as a *hypertext*, which has a non-linear organization, but at the same time needs a linear manifestation when it is interpreted as "instructions for use". The analysis of the linear manifestation of the hypertext requires certain interpolations for an explicit description of co-reference relations established between verbal and

iconic text components, the latter including non-linguistic symbols, colours and typography.

Study (9) by IMRE BÉKÉSI analyses a short extract from L. Tolstoy's *Anna Karenina* (First part, XII, 3-4; Hungarian translation by László Németh). Although the article deals with a *novel*, what interests the author is not so much the narrative aspect of the text, but rather the *argumentative* role of the so-called *key sentence*, which resumes and prefigures the semantic content of a paragraph (or a larger text segment, eventually the whole novel), while the latter gives a detailed explanation of the situation having been referred to cataphorically. Coreferential indexation shows explicitly that the key sentence contains lexical reference to all essential world entities that form co-reference chains in the following text segment. In order to demonstrate the argumentative character of the key sentence and of its relationship with the following text segment, the author also analyses the logical macro-structure of the whole extract organized according to relations of explication-deduction and opposition.

Study (10) also deals with a *literary* text: KATALIN B. FEJES examines a poem, *The Greeting of Thomas Mann* by Attila József, and insists on a special characteristic of this text, which distinguishes it from all the other texts presented in the volume. As it is indicated in its title, the poem depicts an imaginary *communicative situation*, in which the poet speaks on behalf of the audience, "we", and addresses the guest, "you", by a series of verbs in the imperative form, performing the global speech act of "request". The imagined situation of oral communication implies the simultaneous functioning of three channels, auditory, visual and kinetic, each of which is considered here as an autonomous level of co-reference chains. The specificity of coreferential representation in this article is that not only world entities, but even particular acts of request are designated by coreferential indices, since it is those acts, inspired by the situation and the poet's emotions, that form the core of the structure of the poem.

Study (11) by IRMA SZIKSZAI NAGY presents a short *newspaper article*: 'Bike-*party* leader' (*Kurír*, VII [1997], 5, p. 16). The Hungarian title of the article contains a pun on the word "party" (used here in the political sense), and suggests an ironic comment on a small act of charity by a well-known politician. Taking into account the special effect of the tone of the text, the author combines coreferential and stylistic analyses, thus offering a description partly different from that proposed in the orientating article. Apart from a few slight modifications (e.g., a simplified notation) in the representation of co-reference relations — Arabic and Roman numerals for text sentences and for coreferential elements respectively —, some more profound changes are introduced in the analysis, in order to show, in several diagrams, the various textual manifestations of co-reference chains built on dominant lexical units. The author

also studies the distribution and the relative frequency of coreferential elements, both nominal and verbal, forming "clusters", and having an impact on the style of the text.

The last analysis of this volume, Study (12) by GÁBOR TOLCSVAI NAGY, is again concerned with a *literary* text, but a very special one, a so-called "dictionary novel" by Ferenc Temesi, *Dust*, article: 'ideal'. Thus, this passage partly bears the characteristics of an *encyclopaedia entry*, as the one presented in Study (3), except for the personal involvement of its narrator and the slightly mocking tone he uses to speak about his childhood ideal and the disillusion of adulthood. Some of the coreferential phenomena are similar in the two entries, for instance the use of non-linguistic symbols, but in other respects, the analysis discovers notable differences as to possibilities of interpolations and explanations, which, in the case of the novel, seem less evident. This difference is probably due to the fact that the objective information given in the encyclopaedia entry can be set against a background of commonly known frames and schemata, while the novel allows more freedom of interpretation, implying a somewhat less rigorous co-reference indexation.

♦

As a summary of our presentation of various co-reference analyses, we shall briefly outline further research on the subject, as it appears in two, more recent volumes of *Officina Textologica*.

Volume 4 (2000), based on the typological investigations of *Volume 2*, discusses, in 7 articles, issues of a general scope in a theoretical approach. The volume starts with a survey, by GÁBOR TOLCSVAI NAGY, of the international specialist literature in text linguistics / textology exploring problems of co-reference and its relation to anaphor. EDIT DOBI elaborates a complex, semi-formal, semiotic-textological model of Chomskyan inspiration, proposed as a broader framework for the description of co-reference relations, a particular aspect of which, namely the role of thesaurus-like lexico-semantic explications in coreferential analysis, is presented by JÁNOS S. PETŐFI and EDIT DOBI. The last study in the volume, by ANDRÁS KERTÉSZ, examines the problem of 'reference' and 'co-reference' in the language philosophy of Frege, where 'reference' equals the referent (*Bedeutung*) of a linguistic expression, while 'co-reference' is explicated as the identity of referents between two linguistic expressions, whether they have the same sense (*Sinn*) or not. — Beside these theoretical studies, the volume contains concrete analyses of coreferential phenomena, in a highly rhetorical poem (LÓRÁNT BENCZE), and in an extract from *New Testament* (KÁROLY I. BODA—JUDIT PORKOLÁB), the latter article demonstrating a new model for computer-based text processing. Finally, a

practice-oriented study by ZSUZSA BENKES proposes a series of creative exercises for the analytical examination of co-reference relations.

Volume 12 (2005) marks a certain change in research orientation: while previous co-reference analyses worked with Hungarian texts, this time coreferential phenomena are examined in a polyglot contrastive approach, comparing Hungarian with French (3 articles), English (2 articles) and German (1 article). The analyses are either based on linguistic data obtained from various types of corpora, non-literary and literary texts (occasionally a text and its translation[s]) or on examples construed by the authors for the illustration of a coreferential phenomenon. The former method is used by ANDREA CSŰRY in her analysis of dialogue in French and Hungarian literary texts, by ISTVÁN CSŰRY, who uses a large corpus of political debates, also in French and Hungarian, by KÁROLY I. BODA—JUDIT PORKOLÁB in a comparison of a poem by T. S. Eliot and its two Hungarian translations, and by EDIT DOBI, who, in a German-Hungarian contrastive analysis, examines two extracts (with their translations) representing two text types: one written in specialist language, the other being a literary narrative. Finally, two studies investigate particular semantic-grammatical problems: co-reference relations in the conceptual structure of English modal auxiliaries and their Hungarian counterparts, by PÉTER PELYVÁS, and pronominal representation of co-reference in French and Hungarian by SÁNDOR KISS—FRANCISKA SKUTTA.

These three volumes of *Officina Textologica*, covering all main aspects of co-reference relations, will hopefully serve as a starting point for further theoretical research, and might as well be useful for educational purposes.

References

Petőfi S. János 1997. *Egy poliglott szövegnyelvészeti-szövegtani kutatóprogram. Officina Textologica 1*. Debrecen: Kossuth Egyetemi Kiadó. 24–38.

Petőfi S. János 1998. Koreferenciális kifejezések és koreferenciarelációk Példaszöveg: Mt. 9, 9-13. Máté meghívása. In: *Koreferáló elemek – koreferenciarelációk* = Officina Textologica 2. Debrecen: Kossuth Egyetemi Kiadó. 15–31.

Petőfi S. János–Dobi Edit 1998. Utószó. In: *Koreferáló elemek – koreferenciarelációk* = Officina Textologica 2. Debrecen: Kossuth Egyetemi Kiadó. 238–261.

4.
On the Results of the Discussion about the Phenomenon of Linearization of Text Sentences
based on the studies published in the volumes of Officina Textologica[1]
EDIT DOBI

1. Introduction

This study deals with the issue of linearization which is a phenomenon occurring in both sentence grammar and text linguistics. In terms of the theory of science, this statement raises the question as to where is the border between sentence grammar and text linguistics. This is especially interesting in the case of a phenomenon which reflects more or less the same function of language but on different levels, i.e. on sentence level or text level. Now we might as well discuss the theoretical dilemma of distinguishing the two levels; but our main concern here is the overview and assessment of the results of the research concerning the topic of linearization, and, consequently, we intend to deal with the issues of the theory of science only to a certain depth, when they are closely related to our topic.

The primary aim of this study is to provide a review and assessment of the relevant studies published in the Officina Textologica edited by János S. Petőfi, and to present the results, as well as further issues to discuss and problems to solve, of the theoretical discourse and analytical examination, performed in the framework of the polyglot textological-text linguistic research program, concerning the topic of linearization.

The studies to be overviewed were published roughly every year between 1999 and 2003 in the volumes of Officina Textologica. This can be important with a view to the scientific context concerning contemporary international research in textology and cognitive linguistics which has been a major impact on the Hungarian textological research in the last few decades. In his monograph on textology published in 2001 Gábor Tolcsvai Nagy set up a scientific framework of two components as a theoretical background, expressing his basic principles both in a functional linguistic and communication system and in a cognitive linguistic framework (Tolcsvai 2001: 10–11). This approach which connects functionality and cognition and, at the same time, takes into consideration the role of the communicative context in influencing linguistic formulation

increasingly affects the scientific discussion in the volumes of Officina Textologica.

Research on the problem of linearization concerning word order and sentence order is of vital importance. As regards the previous volumes of *Officina Textologica*, within the framework of the polyglot text linguistic-textological research program the topic of linearization has been discussed in the following volumes:

Officina Textologica 3. Towards the analysis of the linear arrangement of text sentence constituents. (Analysis of Hungarian texts.) Szikszainé Nagy Irma (ed.). 1999.

Officina Textologica 6. Towards the analysis of the linear arrangement of text sentence constituents. (Analysis of Hungarian texts. 2: Discussion.) Szikszainé Nagy Irma (ed.). 2002.

Officina Textologica 7. Aspects of contrastive text linguistics. (Linearization: theme-rheme structure.) Petőfi S. János—Szikszainé Nagy Irma (eds.) 2002.

Officina Textologica 9. Aspects of contrastive text linguistics. (Linearization: thematic progression.) Petőfi S. János—Szikszainé Nagy Irma (eds.) 2003.

In the following, I will be presenting the various views and the results of the scholarly discussions on the subject of linearization as follows:

– First, I define the phenomenon with the intention of overviewing all aspects of linearization from the theory of sentence structure to textology. In addition to the general definition of the phenomenon, I provide further contributions to the definition, based on a number of review studies, which are valid only in certain theoretical frameworks.

– Second, I survey the studies on linearization in order to form a clear picture of the results of the theoretical and practical research on the topic. Therefore, I arrange the reviews of the studies according to their content and theoretical background rather than their publishing date.

– Third, after a thorough review and evaluation of each study, I give a summary of the results and conclusions based on the examinations performed in the framework of the research program, in addition, I will suggest some directions for further research.

As regards the structure I intend to follow in the reviews of the studies, I have to remark that

– due to size limitations, the equal treatment of the research activities and results of each author is, as a matter of fact, impossible; so I have to point out the studies that I consider as milestones for the research of linearization that has been carried out during the discussion held in the different volumes of *Officina Textologica*.

– the first and second parts cannot be strictly separated because some review studies express definitive conclusions on certain types of linearization (e.g. FSP, thematic progression).

2. Contributions about the definitive characteristics of the phenomenon of linearization

2.1. About the dominant features of the general definition of linearization

The concept of linearization means, both in the theory of sentence structure and in textology, the arrangement of constituents within the actual language and non-language environment (or context and co-text, according to Petőfi). In a sentence grammar framework, this means the arrangement of words within a sentence unit (clause, simple sentence) and the order of clauses within a complex sentence; in addition, in a text linguistics framework, this means the sequence of sentences. Regardless of the complexity of the language elements used, their linear arrangement is determined by their function in the actual communicative context.

Accordingly, in the chapter about linearization in the first volume of Officina Textologica, János S. Petőfi defines "the meaning related to the linear arrangement of utterances (text sentences) and/or their constituents" as *communicative meaning*, in addition to referential and mental meaning (Petőfi 1997: 38).

Petőfi attaches great importance to communicative meaning in the process of producing, receiving and interpreting the text. This meaning is related to the theme-rheme (or topic-comment) structure of text sentences, and to the thematic progression in the sequence of text sentences.

János S. Petőfi's model of representing text as a complex sign distinguishes between the relational and linear organization of the formal and semantic composition of the actual vs. extended text vehiculum. Relational organization consists of the formal / syntactic, linguistic – semantic and/or thematic relationships between constituents; and linear organization is based on theme-rheme (topic-predicate) structure and thematic progression (Petőfi S. 2002: 51). As regards the aspect of order, in the definition of linearization Petőfi does not differentiate between theme and topic, and rheme and predicate.

2.2. Additional features of the definition of linearization based on the discussion in Officina Textologica

At the present stage of contemporary theory of sentence and textology it is hard to believe that we can define linearization by giving a finite number of

definitive features in a general and acceptable theoretical framework. Rather, we can obtain a more subtle picture of the phenomenon if we analyse in detail the consequences of the different models describing sentence or text, and then try to determine a set of distinctive features to characterize the different phenomena of linearization in general or type by type. Note that, because of the multiplicity of perspectives and approaches of the different schools dealing with the topic, this task might be rather difficult in itself.

At this point let me select some review studies from Officina Textologica which contribute significant additions to the definition of linearization.

In the epilogue of the third volume of *Officina Textologica* (entitled *'Towards the analysis of the linear arrangement of text sentence constituents'*), the editor of the volume, Irma Szikszainé Nagy, summarizing the analyses of the studies, makes general remarks on linearization which are fundamental to the definition of the phenomenon. In addition to the general aspects of the theory of sentences and textology concerning linearization which is well-known in the literature, she calls our attention to the dual nature of the phenomenon.

This duality is basically related to the two terms which are used to describe the phenomenon: linearity *and* linearization. *Linearity* occurs more frequently and has two parallel meanings: (a) the process which results in the structure of text; (b) the result of such a process, the linear (possibly continuous) feature of the language object. According to the author, the first meaning can be associated with the term 'linearization' while the second one with the term 'linearity'. From the viewpoint of the theory of sentences and textology, János S. Petőfi uses the term 'linearization'; while he prefers the term 'linear arrangement' when the theme-rheme structure and the thematic progression is concerned, or when speaking in general about the order of words in sentences or the order of sentences in texts.

When speaking about 'order' in Hungarian linguistics, we must mention the fundamental results of László Deme and Imre Békési. Several authors make reference to the two central concepts of Deme, which are 'megszerkesztettség' (approx. 'structurability'; i.e. the sentence has a certain structure based on grammatical rules) and 'beszerkesztettség' (approx. 'incorporativity'; i.e. further grammatical rules apply when sentences are linked and incorporated into a text). These concepts do not completely fit the approach which attributes a central role to the context in the definition of linearization; nevertheless, from a structural viewpoint, they clearly express the dependency of linguistic structures upon a more complex structure. Deme's functional approach attributes communicative role to the sentence as a whole. However, when describing sentence structure he pays careful attention to the thematic organization of text depending on the given context. In his model for the description of sentence structure the duality

of linear progression and hierarchical organization is based on these principles (Tolcsvai Nagy 2002: 13).

In the sixth volume of *Officina Textologica*, Gábor Tolcsvai Nagy's study entitled *'An outline of linguistic research in Hungary on the functional relationship between sentence and word order'* reviews the major ideas about constituent order, within the scope of a sentence as well as in a broader framework, from the viewpoint of contemporary schools of syntax and textology.

Piroska Kocsány's study entitled *'Questions about theme, rheme and text'* deals with the origin and change of the concepts *theme* and *rheme* „from the dichotomy of theme-rheme to the duality of topic-comment" on the basis of the works of the Prague School. The author explores the duality of a syntactic and communicative approach, and discusses various aspects of the issue of order within the scope of sentences and/or texts.

Gábor Tolcsvai Nagy's and Piroska Kocsány's study, complementing each other's thoughts and ideas, give a comprehensive and summary description of the theoretical background of the issue of linearization. In the following we shall overview the main ideas of these studies.

Focusing on the phenomenon of thematic organization, Gábor Tolcsvai Nagy assesses approaches to the theory of sentence that deal with the question of order. The author provides a number of facts about the definition of the phenomenon of linearization, some of which will be summarized below.

Assessing 'Academic grammar',[2] and the descriptive grammars following and based on it, the author complains that the structural and functional description of word order has been pushed into the background. However, compared to the early structural approach of 'Academic grammar', the new grammatical and structuralist approach of 'Hungarian grammar'[3] to the theory of sentence brings about considerable changes, this, however, does not account for the fact that Keszler's approach to grammar does not at all incorporate the systemic description of word order, which is one of the language-specific phenomena of the theory of sentence, into its theory and above all into its practice; so the functional approach to the description of sentence remains in the background.

As regards thematic organization concerning linearization, the author lists, from the Hungarian researches on the theory of sentence, the functional-based

---

[2] The author refers to a two volume work on the system of contemporary Hungarian (Deme L.; Farkas V.: A mai magyar nyelv rendszere I-II. Budapest: Akadémiai K. 1961.), which has been widely accepted as normative since right after its publication.
[3] The author refers again to an authoritative book on Hungarian grammar (Keszler B. (ed.): Hungarian grammar. Budapest: Nemzeti Tankönyvkiadó, 2000.)

model of László Elekfi and the transformational generative grammar-based model of Katalin É. Kiss; and from the international literature, the researches of Givón, Halliday, Langacker and Danes.

Tolcsvai Nagy considers László Elekfi as one of the main initiator of the functional description of sentence in the last few decades who built up his functional-structuralist theory of sentence on the basis of the results of the Prague School. In this framework Elekfi calls the two functional sentence element theme and 'propositum' in order to distinguish them, firstly, from theme and rheme used by the Prague School, and, secondly, from topic and comment used in the generative description of sentence. As regards the relationship between the two functional elements, „the propositum expresses the essence of utterance with reference to the other element of the sentence, i.e. the theme" (Elekfi 1986: 24, cited by Tolcsvai Nagy 2002: 14). However, this relationship between theme and propositum is not valid for every sentence; but when it is, the place and order of the two functional elements in the sentence are not fixed. Therefore in Elekfi's system the theme is the point of reference, the rules regarding the linear arrangement of constituents are incidental, and the order of the two functional elements can be reversed.

In the framework of transformational generative grammar, the categories topic and comment indicate "it is known" and "it is new", respectively. In the description of the topic-comment structure, the essential generative element is the transformation between the original and the surface structure by which certain elements of the original structure are transferred to topic, comment, or focus position according to certain considerations (Tolcsvai Nagy 2002: 15). Katalin É. Kiss, who, according to Tolcsvai Nagy, has elaborated the most well-known description of the topic-comment categories in the Hungarian generative school, defines the above categories as follows: "Topic is the first nominal and/or adverbial segment of the sentence which is unstressed, while comment is the second structural segment of the sentence which follows the topic, begins with the part of the sentence being stressed, and contains the verb" (É. Kiss 1983: 16, cited by Tolcsvai Nagy 2002: 15). Compared to that of Elekfi, this definition clearly provides much more systemic statements concerning the linear arrangement of constituents.

In her later works Katalin É. Kiss specified the meaning of the category 'focus' by adding that focus expresses exclusive identification, and improved the description of topic-comment categories with special regard to the structure of sentence as follows:
– the structure of a prototypical sentence follows a topic-comment sequence;
– that particular topic-comment structure when the comment starts with a focus is of great importance;

– there are certain structures which have no topic, only comment (or focus followed by the other parts of the comment);

– the two main parts (i.e. the topic and the comment) are not interchangeable; the same sentence constituent in different places of the sentence can be topic or (part of the) comment (É. Kiss 1995: 15, cited by Tolcsvai Nagy 2002: 16).

The adequate description of the problem of linearization can possibly be conceived by combining the results of Elekfi and É. Kiss. The same observation can be made in the relevant international research; both the functional approach of Halliday and Givón, and the cognitive approach of Langacker, although in different theoretical frameworks, deal with the structure of sentence from a semantic viewpoint. They all agree that the description of structure and that of function should be combined (Tolcsvai Nagy 2002: 17).

Halliday's functional grammar (Halliday 1985) introduced for the first time two methods for the description of sentence that can be used together by completing each other: on the one hand, a method "that is formal, identifying sentence constituents directly, and describing phrases"; on the other hand a method "that is functional, classifying sentence constituents, and describing parts of a sentence". The description of the theme-rheme structure of the sentence, according to the Prague School, can then be based on the twofold system of the two above-mentioned methods (ibid.).

In his theory, Givón examines thematic continuity in a sequence of sentences (which can be either a paragraph or a text) which involves studying action and topic continuity. According to Givón, these aspects together determine a larger text unit, e.g. a paragraph.

Langacker's cognitive model deals with the understanding of a sentence and tries to formalize the process of sentence comprehension. One of the clues for this effort can be the interpretation of the topic or topics in the sentence which in itself does not include any rule as to the linear arrangement of the sentence. However, the word order in the sentence can be explained by a so-called 'cognitive path' paradigm based on the scheme 'source – path – target' where the interpretation process can go from the subject (i.e. source) to the direct object (i.e. target) or vice versa. The direction of the process depends on which meaning should be foregrounded (cf. Tolcsvai Nagy 2002: 20).

It can be easily accepted that there are different rules and relationships to be considered within a sentence or a sequence of sentences (or text) concerning the description of the linear arrangement of constituents. Another approach to the question might be that the ratio between grammatical rules and contextual (or situational) effects is different when speaking about a sentence or a text. According to Tolcsvai Nagy, the fact that research in Hungary on thematic progression in the text or in a given part of the text lags behind international research can be explained by the focus on the sentence in the Hungarian

generative research tradition.. The Hungarian theoretical frameworks are based on two models, i.e. Danes' model of thematic progression and É. Kiss' model of the thematic organization of text, respectively (ibid.).

The most concrete conclusion Tolcsvai Nagy draws about the current situation of the description of sentence and text explored in his study is that "the functional sentence perpective is discourse-dependent, therefore the immediate context and situation needs to be modelled for the description to be valid. Consequently, the attempts of textology support, from among the previous theoretical and descriptive approaches, the functional and cognitive approaches, at the same time synthesizing the results of the formal generative theory." (Tolcsvai Nagy 2002: 21) The slight but noticeable trend changes in the theory of science toward cognitive approaches have verified Gábor Tolcsvai Nagy's observations about dealing with texts, which he expressed more than ten years ago.

Some of the analytical studies published in *Officina Textologica* represent the same tendency toward the examination of the phenomenon of linearization. Exploring systematically the contextual-logical relationships in text, Károly I. Boda and Judit Porkoláb provide a formal description of the theme-rheme relationships and the thematic progression within particular literary texts. Applying the so-called co-reference analysis elaborated by János S. Petőfi, they intend to analyse and formally describe the deeper structure of the meaning of text. (Boda–Porkoláb 2002 and 2003). Other studies compare the linear arrangement of text sentence constituents in Hungarian texts with that of foreign (i.e. non-Hungarian) language texts taking into consideration the cognitive aspects of texts (Csűry A 2002 and 2003; Pelyvás 2002 and 2003; Skutta 2002 and 2003; Kiss 2002).

In her study mentioned above, Piroska Kocsány analyses the dichotomy of theme-rheme and topic-comment on the basis of research by the Prague School concerning sentence and text linguistics. With respect to text, mostly in Danes' research "the theme gets a new accent" which leads to the concept of thematic progression that can be applied to the examination of texts. According to Piroska Kocsány, "in this framework the concept of theme and rheme puts a broader perspective on the different domains of meaning which can be (or to be) interpreted in various ways".

Piroska Kocsány's train of thought leads to interesting implications for the textological aspects of the theme-rheme phenomenon . Starting from the fact that the syntactic structure of Hungarian sentences is highly constrained, she tries to thoroughly explore and analyse the fundamental issue of whether it is necessary to introduce the theme-rheme operation which can be hard to define, or we should stick to the trichotomy of topic-focus-comment which can be used according to a well-functioning system of rules for describing sentences and can

also serve as a reference point in the description of texts. However, the author points out four areas where the trichotomy of topic-focus-comment, which functions perfectly in the framework of the theory of sentences, cannot be used or can only be used under certain conditions. These areas are as follows: prosody, the question of how to express the intention of the speaker, the question of conveying cognition and style, and the information structure established in the process of producing and processing text (Kocsány 2002: 13). In the remainder of her study the author examines these four areas of language production with a view to answering the question of how best they can be characterized: along the lines of the theme-rheme dichotomy or the topic-focus-comment trichotomy.

As one might guess, the answer is quite difficult and not necessarily definite; however, it is significant from the viewpoint of research on the topic. The author arrives at the conclusion that "for systemic linguistics, the use of the concept of 'topic-comment' is safer than the vague concept of 'theme-rheme'. For the most part, this statement holds true, and is especially valid for the explanation of (phycho)linguistic questions related to word order or the understanding of sentence in the case of the Hungarian language. Using the 'topic-comment' construction can also be useful for prosodic studies even if here we should take certain special cases into consideration the explanation of which are far beyond the limits of systemic linguistics. As a consequence, the 'topic-comment' construction, after it has been precisely elaborated, can help us perform certain prosodic studies as well as examine the structure of discourse. However, the main function of the 'topic-comment' dichotomy is to structure the text according to contextual factors. As a result, we should go beyond the traditional linguistic concept of 'theme-rheme'. […] Since the extended interpretation of the 'topic-comment' construction (whether functional or cognitive-based) is focussed either on the cognitive units of the process of producing and processing text or the identification of mental operations, and not merely on language-level facts, it obviously goes outside the functional framework of linguistics." (Kocsány 2002: 17)

3. About the lessons that can be drawn from the discussion held in the framework of the polyglot text linguistic – textological research program

One definite advantage of the he polyglot text linguistic – textological research program is the existence of systematic analyses that methodologically complement the theoretical research. In the first volume of Officina Textologica János S. Petőfi, outlining the basic aims of the research program, emphasizes two basic features:

– on the one hand, the program is to be polyglot because we are convinced that the textological – text linguistic characteristics of (the texts of) a specific

language cannot be explored in-depth if we do not examine them from the viewpoint of at least another language;

– on the other hand, the program is to be integrated as far as possible; we would like to encourage scholars from different language backgrounds and interests to study the textological – text linguistic phenomena in question (also) from more or less identical — or at least an explicitly comparable — viewpoint. (Petőfi 1997: 7–8)

In harmony with these principles, first, some linguists participating in the scientific discussion (concerning also the topic of linearization) compare the overall characteristic of Hungarian texts with English, German, French, Russian and Italic texts and other language attributes; and second, every research task is part of a well designed and elaborated research project the research methodology of which enables scholars from different theoretical background to cooperate with each other fully.

The studies of *Officia Textologica* dealing with the phenomenon of linearization can be broadly classified into four categories:

1) I have thoroughly reviewed in the foregoing discussion the review articles that have dealt with the theoretical aspects of the 'theme-rheme' structure.

2) Other studies examine word and sentence order in the Hungarian language in comparison with other (i.e. not Hungarian) languages.

3) Further studies analyse the phenomenon of linearization on the basis of a given model or method, and draw conclusions from empirical results.

4) The fourth set of studies extend their investigations concerning the phenomenon of linearization to the more general aspects of text. In these cases the overall assessment of the phenomenon of thematic progression is necessarily beyond the scope of linguistics, although it applies the lessons drawn from the investigations of the structure of sentence.

In the following I will focus on the three remaining groups of studies.

3.1 On the results of the contrastive text linguistic studies

The studies of the seventh volume of *Officina Textologica* (*Aspects of contrastive text linguistics*, ed. by János S. Petőfi and Irma Szikszainé Nagy) compare the theme-rheme organization of Hungarian sentences to that of English and French sentences.

Péter Pelyvás studies the theme-rheme organization of English sentences from a cognitive grammatical perspective. He characterizes the relation of the phenomenon of theme-rheme structure to word order and grammatical functions compared with the structural features of Hungarian sentences. The contrastive study of English and Hungarian sentence structures seems to be useful because of the significant difference between the two languages in this respect. A slightly

simplified explanation of the rigid word order in English is that the grammatical function of words is determined by their position in the sentence. On the other hand, the so-called free word order in Hungarian is motivated by the theme-rheme structure because the grammatical functions are determined by grammatical morphemes and not the structural position of words. This fact explains the essential difference in the description of sentence structure between Hungarian and English; in the Hungarian language, the exploration of the theme-rheme structural representation is preferred to the subject-predicate analysis of sentence structure, while the description of the English sentence structure is based on grammatical functions. Analysing certain type of structures (e.g. passive structures, complex transitive predicates, existential sentences, etc.) Pelyvás looks for an answer to the question of "whether the description of English sentence based on the grammatical functions can be replaced with a description which is based on the communicative aspects to a greater extent" (Pelyvás 2002: 20).

The analyses lead to interesting consequences. According to the author, one of the main conclusions is that "the theme-rheme organization is a very important part of the process of building, or being integrated into, a cognitive model" (Pelyvás 2002: 32). In addition, he suggests that we reconsider the traditional theorem that in languages with rigid word order "there is a serious tension between the communicative aspects of the theme-rheme structure and the rules governing the arrangement of grammatical functions within the sentence" because, for example in the English languagethere are only a few structures that can be used to ease this "tension". Moreover, the author holds that the traditional notion of the English language having rigid word order needs to be reassessed because certain communicative contexts enable the creation of non-prototypical structures (Pelyvás 2002: 33). In the author's opinion, holistic cognitive grammar is suitable to describe sentence-level, and preferably text-level, phenomena.

Sándor Kiss and Franciska Skutta study the same tension between the communicative and grammatical aspects of sentence structure in French that Péter Pelyvás examined in English sentences, as we have mentioned before.

The title of Sándor Kiss's study, *"Parts of speech and theme-rheme structure: the conflict of two structures in French"*, clearly expresses the contrast between grammatical function and communication structure. The author's starting point is as follows: "French is usually considered as a language having rather rigid word order." (Kiss 2002: 35) This language feature raises the question of the conflict between the logical-functional structure of sentence and theme-rheme structure. The author mentions two opposite schools in French linguistic science which attempt to handle this conflict. The first approach is the functional view of traditional grammar which considers sentence as "a unit of

functions to be interpreted logically" and takes functional sentence perspective into account only when analysing specially or loosely constructed sentences. The other approach "takes into account the starting point of the message and the aim of the utterance" in the first place (Kiss 2002: 35). Sándor Kiss studies the two views in one and the same framework and examines "the kind of solutions in the theory of sentence that can ensure the preferred handling of the starting and the end point of the message in French presuming that the language user remains within the broad scope of the logical-functional rules concerning the construction of sentences." The author focuses on three phenomena: "putting the rheme on focus position; the structural position or anteposition of the starting point, i.e. the theme; the constructions that prepare and practically preannounce the rheme (Kiss 2002: 36)." As a final conclusion of his analyses, Sándor Kiss draws the lesson that "there is a one-sided relationship between the logical-functional and communicative organization of sentences; the latter can always be reconstructed in the message, but it does not necessarily determine the former".

Franciska Skutta's study examines the thematic organization of complex sentences in French with special reference to the order of clauses. The examination is based on the structural typology of complex sentences and pays special attention to the relationship between clauses and the conjunction expressing this relationship. The author illustrates functional sentence perspective analysing different types of the structure of the complex sentence by an example each. Her analyses confirm the otherwise well-known fact that subordination is essentially a grammatical phenomenon, but the logical relations of coordination lead to the field of textology. According to the author, this difference appears in the theme-rheme structure of the clauses of complex sentences because "determining the theme and pointing out the rheme can in the strict sense be accomplished only in subordination, and the structure can be accurately described by corresponding rules. [...] in the contrastive study of French and Hungarian (or other languages) we might find that subordination structures have more language specific features, and therefore greater differences between languages than coordination which is based on logical relations in the first place." (Skutta 2002: 64)

In her study entitled '*The correspondence of Hungarian text sentences having focus in French translation: possible syntactic structures and tendencies*' Andrea Csűry (Nagy) presents a comparative study of French and Hungarian. Limiting the scope of the study, she deals only with positive affirmative text sentences having a constituent which can be interpreted as a focus. The author examines the question of how we can express and determine the focus in the light of comparing the structure of Hungarian sentences with that of their French translations. Using a corpus of text sentences containing verbs with prefix the

separation of which makes it easier to identify the position of focus, the author wants to find "the basic elements in French, considering that it is a rigid word order language, that correspond in translation to the parts of the Hungarian sentences in focus position" (Csűry A. 2002: 65). Analysing the examples she draws the conclusion that "the French language expresses focus essentially in two ways, i.e. with the emphatic syntactic structure and with stress, respectively" (Csűry A. 2002: 73).

The central issue of István Csűry's research is about the behaviour of connectors. In his study entitled '*The syntax of the thematic structure and the connectors in the French and Hungarian language*' he examines the existence of the relationship between the position of connectors in text sentences (for those connectors that do not have a fixed position) and the tematic structure of text sentences. As regards the issues examined, there are some similarities in French and in Hungarian: on the one hand, "in French, the syntactic rules for adverbial connectors cannot be described without taking into consideration the thematic structure" (Csűry I. 2002: 85–86); on the other hand, "in Hungarian, we should take into account the status of clauses within the co-text, the way that the connection is realized, and the thematic structure, in order to describe the syntax of adverbial connectors, even if the role of these factors is different in the case of various lexical units." (Csűry I. 2002: 89) According to the author, "the thematic structure of the text sentence, and its aspects concerning word order, should be examined in parallel with the thematic organization of the co-text [...] to answer the syntactic issues discussed here we should first analyse the ways in which connection can, in the given case, be realized, and not necessarily the theme-rheme structure within the sentence" (Csűry I.: 2002: 89).

The four remaining studies of the seventh volume of *Officina Textologica* make valuable, and mainly theoretical, additions to the contrastive and analytical approaches described above.

Károly I. Boda and Judit Porkoláb's main concern is to examine poetic texts. In their study entitled '*Examination of the theme-rheme structure of a selected poetic text using co-reference analysis*' the authors present a step-by-step methodology for analysing a selected poetic text which, in this case, is a poem by Milán Füst. The main conclusion of the authors is that "co-referential analysis can be effectively applied to explore the theme-rheme structure of a given text" (Boda–Porkoláb 2002: 93).

Edit Dobi's formulation-based approach to the organization of text sentences is mainly based on structural principles, including semantic aspects of sentence structure, as well. In her study entitled '*To the analysis of the division of text sentences on the basis of given and new information*' she describes, starting from the generative description of sentences, the role of pragmatics in the

representation of systemic and text sentences including the representation of the linear arrangement of constituents and the theme-rheme structure.

Zoltán Szabó's main concern is to explore the status of comparative stylistics in the system of textological disciplines, just as he has accomplished it so far in the case of other disciplines such as contrastive linguistics, translation studies, and comparative literature studies. In his study entitled '*Linearity and comparative stylistics*' he examines the phenomenon of linearization from a stylistic standpoint providing detailed analyses of several examples. These examinations are considered to be significant and justified because stylistics, and especially comparative stylistics has not yet taken enough attention to the phenomenon of linear arrangement.

Olga Murvai's study entitled '*Theme-rheme structure and the question of translation from a textological viewpoint*' deals with the manifestations of the theme-rheme structures of the target language in translations. The author extends the issue of linearity to a textual phenomenon. Her starting point is the topic-focus dichotomy based on the theory of Sanford and Garrod. The author distinguishes between explicit and implicit foci as follows: "explicit [focus] concentrates on a specific entity of the text whereas implicit [focus] concentrates on a script or the background knowledge which characterizes the corresponding entity" (Murvai 2002: 129). Her basic question is "whether the distinction between implicit and explicit foci can be efficiently used in translation" (ibid.). The main conclusion of the author's very illuminating experimental analysis is that the above dichotomy presents an informative basis for translation.

In the final study of the volume entitled '*»This« functioning as cataphora in the English language*', Olga Bársony expounds the idea that the use of 'this' under discussion (instead of using the indefinite article 'a(n)') sends a specific message to the receiver, i.e. that the speaker is inclined (or might definitely want) to provide extra information about the topic, or discuss it. Concerning this phenomenon, the author deals only with its occurrence in the English language and does not compare its parallel to that of the Hungarian language.

3.2 The practical aspect of the text linguistic discussion

As reflected in the previous studies mentioned before, the polyglot research program involves both theoretical discussions and analytical approaches. As we have seen before, the detailed outline of the phenomenon of linearization required theoretical and integrated approaches in the first place, as well as analytical works written as a kind of justification of a particular theoretical hypothesis. In parallel with it, it might be worthwhile, and perhaps interesting as well, to have a good look at those methodological starting points or analytical techniques the application of which can shed light on the various features of

sequentality. These techniques are represented by particular authors, or groups of authors, in the research program which edits and manages the periodical *'Officina Textologica'*. The approaches followed by the authors necessarily correspond to each other in that the target of the analyses should be considered as a syntactic and semantic structure influenced by its context.

The studies which are intended to carry out an analysis can be basically divided into two groups: the first one contains studies the authors of which have undertaken a given task applying a prescribed methodological framework; whilst the second one contains studies whose authors examine a specific phenomenon that has been established before, applying individual and unique (or at least not uniform) methods.

The studies from the first group can be found in the third volume of *Officina Textologica* entitled *'Towards the analysis of the linear arrangement of text sentence constituents. (Analysis of Hungarian texts.)'*. Based on the works of János S. Petőfi and Zsuzsa Benkes, we call the analytical approach which the authors of these studies follow **creative-productive text processing**. In addition to the fundamental role of the system of grammatical rules concerning the word order in sentences, or concerning the representation of linguistic structures to be more general, the creative-productive approach emphasizes the significance of intuition in the process of producing and processing a sentence or text.

All the authors of the volume have undertaken a (specific variation of a) task of analysing a given part of text assigned to them by János S. Petőfi. In the foreword of the volume, Petőfi says about the research:

"For the target of the analysis, we have selected three different parts of the same literary text, namely the first five paragraphs of 'The lynx', a short story by Lajos Áprily. The selections to be analysed were as follows: (I) the first clause of a compound text sentence; (II) a complete but simple text sentence; (III) a compound text sentence which contains five syntactically independent clauses. The authors have been asked to analyse the possible linear arrangements of the above selections.

We have made special preparations for the analyses. In each case the selection (either a clause or a text sentence) has been deleted from the original text and we gave all the possible linear arrangements of its constituents without telling the authors which was the actual arrangement that occurred in the given context. Giving the authors a creative exercise, the main task of the analyses was to decide which arrangements in each case can be considered acceptable and which cannot.

We have asked three persons for each type of selections; two of them were asked to actually perform the analysis, and one to make his or her students of a seminar do it, and then summarize and evaluate their results." (Petőfi 2002: 11)

I am convinced that the detailed description of the task was worth mentioning because it determined the method the authors employed. Although the authors of the volume have undertaken an analytical task, they also deal, extensively or meditatively, with theoretical issues; it is because of the nature of the matter discussed which is rather complex and not very well-defined. In the epilogue of the volume, Irma Szikszainé Nagy gives a comprehensive account on the results of the analyses.

In her study entitled '*Thematic progression and linear arrangement*', Franciska Skutta analyses not only the given variations in the order of words in selection (I) but also examines further possible variations. She evaluates the various forms of linear arrangement dynamically, according to the reception of the content. In his study entitled '*The role of prosody in the formation of linear arrangement*', Imre Wacha handles the assigned task uniquely, examining the subject on the basis of prosody which is the main concern of his research. In her study entitled '*Examination of linear surface structures on the basis of five paragraphs selected from the short story »The Lynx« by Lajos Áprily*', Edit Dobi summarizes the results of an experiment conducted by a group of students. Her analysis follows a generative approach taking into consideration the textological aspect that both the given context and the factors of the current communication situation have impact on the linear arrangement of the constituents of a text sentence. The title of Piroska Kocsány's study is '*Variations in the order of words in texts: which are »authentic«?*'. Taking an analytical approach, she focuses on the context that follows the sentence analysed which is rather unique because most analysts prefer the context that precedes the sentence. She concentrates on how the prosodic features of the sentence overwrite the rules of the theory of sentence structure. In her study entitled '*Creation of a fairly acceptable vehiculum for a piece of prose*', Eszter Szikoráné Kovács declares that we should evaluate the information value and stress relationships in each word-order variation in order to select the acceptable one(s). In his study entitled '*A creative-productive exercise in assessing the formation of a selected verbal text*', László Vass summarizes the analysis of a college strudent group. He points out that the students almost invariably selected the most acceptable variation directly, but when trying to select the least acceptable one the results were more random. In her study entitled '*A creative approach to the possible linear arrangements of text sentence constituents in clauses (or simple sentences)*', Ágnes Domonkosi deals with the manifestation of linearization in text including the characterization of thematic progression. On text level the linear arrangement of logically related sentence constituents (supposing that they are connected without conjunctions and do not include co-referential elements that have an effect on the sequence of the constituents) is far less rigid than the word order in the sentence. According to the author, the acceptability of the different variations

is determined by the direction of the description of the content influenced by various factors (in this case the suggested description of 'the lynx' goes from a general picture to specific details), as well as the rhythmic and syntactic features. In his study entitled '*Analysis of the possible linear arrangement of independent clauses*', János S. Petőfi analyses the factors that determine the order of syntactically independent clauses in a compound text sentence. The author introduces a system called "triple filter" which is built on three pillars: first, he analyses the logical structure of each clause and tries to formulate rules on the order of clauses; second, he examines the rhythmic structure of the possible linear arrangements of the clauses; third, he studies the effect of perception on the possible linear arrangements. In her study entitled '*The analysis of communicative meaning in the linear arrangement of the constituents of a text sentence*', Katalin B. Fejes summarizes the work of a group of students. She reveals that the students' choices have been influenced by logical-semantic factors (e.g. the whole-part relationship between the elements of the description, the order of the general-important and specific-unimportant elements, etc.) and syntactic-rhythmic considerations (e.g. the length of a sentence constituent, the occurrence of the verb as a predicate in a sentence, rhythm, rhymes, etc.).

As I mentioned before, in the epilogue of the volume Irma Szikszainé Nagy presents a summative evaluation of the results of the authors' analyses. She systematizes the methodological and theoretical conclusions and emphasizes the deficiencies to be overcome in the future investigation of the subject. She voices her opinion that a thorough and in-depth examination of the subject must be carried out.

The discussion of the third volume of *Officina Textologica* is included in the sixth volume. The first study of the sixth volume is written by Gábor Tolcsvai Nagy, his conclusions were presented in detail when we were discussing the definition of the phenomenon of linearization. The other studies provide an intriguing extension of the creative-productive approach to text. Except for Zsuzsa Benkes, who in her study carries out a "classical" creative-productive text analysis, the authors complete the approach invented by Petőfi and Benkes with their own views and ideas. (Under the methodological classification scheme mentioned earlier in this study, these studies belong to the second group where the creative-productive approach they adopt is realized in a unique way.)

The authors of the sixth volume continue the analysis on the selected passage by Áprily now making use of the results of the previous studies.

The title of Edit Kádár's study is '*Bábirkó-style wrestling*' which is itself a witty and thought-provoking pun expressing the frequent struggle that linguists, doing their analysis, carry on with the text in order to find some general rules concerning its organization. Within a slightly modified framework of semiotic textology, the author distinguishes between the grammar of the speech product

and speech production grammar (which needs to be elaborated, the author adds), emphasizing that in the speech process text is a product and production at the same time. As regards the first approach (i.e. considering text as a speech product), the author analyses the different variations on the basis of lexico-grammatical and phonetical-prosodic features, and the communicative or thematic organization of text. As regards the second approach, the author attempts to evaluate the communicative intention behind the different variations.

János S. Petőfi's study entitled 'An analytical approach to the linearized structure of texts' outlines a typology of the organization of text worked out by the author, and provides a step-by-step introduction to the analysis intended to result in an explicit and formal description of text. Petőfi's genuine and brilliant idea is that two different types of organization of text can be distinguished, namely the relational and linearized organization. By analysing these organization types we can describe the hierarchical and linear structure of a given text, as presented by the author projecting the system of organization types onto the text analysed. Finally, he summarizes and raises questions for future research.

In her study entitled 'A creative approach to the linearized structure of texts', Zsuzsa Benkes reviews the analytical methods which she has developed together with János S. Petőfi in order that, as the author says, our systemic linguistic, text linguistic, and textological knowledge, suppositions, expectations and our knowledge about the world, most of which are usually latent, will be revealed in the analytical processing of text. (Benkes 2002: 71) The author describes two fundamental approaches, namely the creative-productive and the kaleidoscopic approach to text, both involving a corresponding method. The first one can be either constructive or selective, both aiming at a given text to be worked on. The aim of the second one is to process a given text according to certain aspects. The author analyses several examples which clearly demonstrate that the described methods can greatly contribute to a better understanding "how text works", taking either language or other aspects into consideration.

In his study entitled 'Management of distance relationships in a totally lexical grammar', Gábor Alberti examines the problem of linearization in a narrower framework where the question of word order within a sentence belongs to the domain of generative grammar (and semantics). His theoretical framework is the so-called GASG grammar (Generative / Gereralized Argument Structure Grammar) which states that "the information which is required for the building of a sentence structure from words comes more and more from the lexical description[4] of the (suffixed) words, and not from the general syntactic rules."

---

[4] i.e. the description of all the words that might occur in the (potential) grammatical context of each suffixed word (ibid.)

(Alberti 2002: 89) The author demonstrates his ideas analysing two sentence examples.

As we have seen so far, the sixth volume of *Officina Textologica* discussed and analysed further the questions that were raised by the authors of the third volume. As a result, in the sixth volume some new and genuine analytical methods have been invented.

3.3 Linearization in a broader framework. Research on thematic progression

The ninth volume of *Officina Textologica* is about thematic progression. There are eight studies in the volume, all of which deal with the assigned topic of the volume, either following a theoretical perspective, or examining selected texts and drawing inferences from the experiences of the examination.

In her study entitled '*»New« types of progression and two-sided rheme structure*', Andrea Csűry (Nagy) deals with the thematic organization of French-language newspaper articles. In the analysed French texts the author observes a particularly economical construction[5] which is characteristic of them. She calls this construction a two-sided rheme structure. (Csűry A: 2003: 16)

In his study entitled '*Thematic progression and its absence in English psychotic narrations*', Péter Pelyvás deals with a specific text type, namely narrative texts created by psychotic patients whose task was to give an account of the events seen in a short movie of five scenes.

In her study entitled '*What a fine piece of argumentation! The caricature of thematic progression in Molière's plays*', Franciska Skutta analyses selected passages from two of Molière's plays. The author focusses on the so-called argumentative sequence of the selected passages and draws the conclusion that the distortions of this structure in Molière's texts are caused by "breaks" in the process of thematic progression. (Skutta 2003: 33)

In their study entitled '*Examination of thematic progression in selected texts using co-reference analysis*', the authors present a complex analytical apparatus which has been elaborated in order to analyse, using a formal and algorithmic approach, the co-reference structure and thematic progression of texts. They demonstrate their conception by analysing a selected passage from a poem by Miklós Radnóti. The overall aim of their research is to develop a general tool for the description of texts "which will enable, in a language-independent way, to

---

[5] In this construction the thematic progression does not follow the usual, „standard" theme-rheme order; instead, the amount of information related to the theme is doubled by an additional rhematic unit which precedes the theme (ibid.).

represent and examine selected formal and material elements of the analysed texts in a knowledge-based information system" (Boda–Porkoláb 2003: 45).

In his study entitled '*Progressive dialogue? (on the basis of selected texts by Molière)*', Sándor Kiss interprets the meaning of thematic progression in a more general way compared to the usual interpretations. Analysing selected dialogues, the author examines the breaks in the process of thematic progression taking careful attention to the relationship between the characteristics of the analysed text and the events of the represented world of the fiction. (Kiss 2003: 59)

In his essay entitled '*Connectors and thematic progression*', István Csűry raises and examines the question "whether it is conceivable to draw general conclusions from the presence of connectors about the thematic organization of their context" (Csűry 2003: 65). After analysing eight passages from selected texts, the author declares that there *is* connection between the connectors and the thematic structure of the cotext. On the one hand, the connectors, which are essential functional elements of the organization of text, might never be ignored when analysing thematic relationships; on the other hand, the thematic structure plays a key role in interpreting the semantic-pragmatic relationship identified by the connectors.

In his essay '*Thematic progression in the context of the history of style*', Zoltán Szabó, who is a leading scholar in the field of stylistics, demonstrates that the typology of the different trends of style can be systematically explored by analysing thematic progression in carefully chosen examples. He draws the conclusion that simple and complex trends of style can be clearly distinguished on the basis of the different types of thematic progression observed in them. (Szabó 2003: 88)

Following a semiotic textological approach, János S. Petőfi characterizes the different forms of text organization in a unique way, one pillar of which is the thorough examination of meaning. In his study entitled '*Different forms of text organization and the three facets of meaning*', the author studies the relationship between the three facets of meaning he distinguishes (i.e. the conceptual, co-referential, and communicative components of meaning) and the different forms of text organization, analysing a selected poem by László Kálnoky. From this point of view, he also examines the theme-rheme structure and the thematic progression in the poem.

In the foreword of the ninth volume János S. Petőfi, the chief editor of the periodical *Officina Textologica*, summarizes the research work presented in the volume as follows:

"The relevance of the contributions, in addition to their individual merits and the fact that they employ a variety of theoretical approaches and analyse different types of texts, is further enhanced by the fact that all the authors

have participated in all *Officina Textologica* conferences organized so far, and have in this way contributed to the creation and maintenance of a forum for continuous professional discussion." (Petőfi 2003: 8)

4. Summary

Thanks to the organization of research activity by János S. Petőfi and the participating researchers who have been ready for the mutual exchange of their views and ideas, the *Officina Textologica* has become a unique textological-text linguistic forum for the presentation, analysis and theoretical discussion of specific problems and the publication of the results. The studies presented here, as it clearly appears from their results, have substantially contributed to the definition and discussion of the problem of linearization and, although there is a lot of research being conducted in the field of textology on theme-rheme organization and thematic progression nowadays, the authors of the studies raise several interesting and important questions for contemporary textological research.

List of the studies about the problem of linearization published in the volumes of Officina Textologica

Alberti Gábor: Távolsági kapcsolatok kezelése egy totálisan lexikális grammatikában. [Management of distance relationships in a totally lexical grammar.] In: Szikszainé Nagy Irma (Szerk.) *Officina Textologica 6.* Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. (Magyar nyelvű szövegek elemzése. Diszkusszió). Debrecen, 2002. 88–112.

Bársony Olga A *this* mint a katafora eszköze az angol nyelvben. ['This' functioning as cataphora in the English language.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 143–149.

Benkes Zsuzsa: A szövegek linearizált szerkezetének kreatív megközelítése. [A creative approach to the linearized structure of texts.] In: Szikszainé Nagy Irma (Szerk.) *Officina Textologica 6.* Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. (Magyar nyelvű szövegek elemzése. Diszkusszió). Debrecen, 2002. 69–87.

B. Fejes Katalin: A kommunikatív jelentés vizsgálata egy szövegmondat összetevőinek lineáris elrendezésében. [The analysis of communicative meaning in the linear arrangement of the constituents of a text sentence.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-*

összetevők lehetséges lineáris elrendezéseinek elemzéséhez (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 120–131.

Boda I. Károly—Porkoláb Judit: Téma–réma kapcsolatok vizsgálata egy kiválasztott versszövegben korreferenciaelemzés segítségével. [Examination of the theme-rheme structure of a selected poetic text using co-reference analysis.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 93–112.

Boda I. Károly—Bodáné Porkoláb Judit: A tematikus progresszió vizsgálata kiválasztott szövegekben korreferencia-elemzés segítségével. [Examination of thematic progression in selected texts using co-reference analysis.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 45–58.

Csűry Andrea: Fókuszt tartalmazó magyar szövegmondatok megfeleltetése francia fordításban: szintaktikai lehetőségek és tendenciák. [The correspondence of Hungarian text sentences having focus in French translation: possible syntactic structures and tendencies.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 65–74.

Csűry Andrea: „Új" progressziótípusok és kétoldali rémaszerkezet. ['New' types of progression and two-sided rheme structure.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 9–16.

Csűry István: A tematikus struktúra és a konnektorok szintaxisa a franciában és a magyarban. [The syntax of the thematic structure and the connectors in the French and Hungarian language.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 75–92.

Csűry István: Konnektorok és tematikus progresszió. [Connectors and thematic progression.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 65–80.

Dobi Edit: A lineáris felszíni struktúrák vizsgálata Áprily Lajos *A hiúz* című elbeszélésének első öt bekezdése alapján. [Examination of linear surface structures on the basis of five paragraphs selected from the short story 'The Lynx' by Lajos Áprily.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 48–59.

Dobi Edit: A szövegmondatok adott és új információ szerinti tagolásának elemzéséhez. [To the analysis of the division of text sentences on the basis of given and new information.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 113–122.

Domonkosi Ágnes: Tagmondatnyi szövegmondat-összetevők lineáris elrendezhetőségének megközelítése. [A creative approach to the possible linear arrangements of text sentence constituents in clauses (or simple sentences).] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 3. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 86–94.

Kádár Edit: (Bá)birkózás. [Bábirkó-style wrestling.] In: Szikszainé Nagy Irma (Szerk.) *Officina Textologica 6.* Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. (Magyar nyelvű szövegek elemzése. Diszkusszió). Debrecen, 2002. 25–49.

Kárpáti Eszter: Megjegyzések a szöveg fogalmához. [Comments on the concept of text.] In: Szikszainé Nagy Irma (Szerk.) *Officina Textologica 6.* Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. (Magyar nyelvű szövegek elemzése. Diszkusszió). Debrecen, 2002. 113–122.

Kiss Sándor: Mondatrészek és téma–réma tagolás: két struktúra konfliktusa a franciában. [Parts of speech and theme-rheme structure: the conflict of two structures in French.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 35–50.

Kiss Sándor: Előrehaladó párbeszéd? (Molière-szövegek alapján). [Progressive dialogue? (on the basis of selected texts by Molière).] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 59–64.

Kocsány Piroska: Szórendváltozatok a szövegben: melyik az „igazi"? [Variations in the order of words in texts: which are 'authentic'?] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 62–66.

Kocsány Piroska: Téma, réma és szöveg: kérdések. [Questions about theme, rheme and text.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 9–18.

Murvai Olga: Téma–réma tagolás és a fordítás kérdése szövegtani nézőpontból. [Theme-rheme structure and the question of translation from a

textological viewpoint.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 129–140.

Pelyvás Péter: Szórend, grammatikai funkciók és téma–réma szerkezet a kognitív grammatikában (Kognitív alapú angol–magyar összevetés). [Word order, grammatical functions and theme-rheme structure in cognitive grammar (A comparative study of the English and Hungarian languages following a cognitive approach.)] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 19–34.

Pelyvás Péter: A tematikus progresszió és annak hiánya angol nyelvű pszichotikus szövegekben. [Thematic progression and its absence in English psychotic narrations.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 17–32.

Petőfi S. János: Egy poliglott szövegnyelvészeti-szövegtani kutatóprogram. [A polyglot textological-text linguistic research program.] Debrecen, 1997.

Petőfi S. János: Előszó. [Foreword.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 7–11.

Petőfi S. János: Mellérendelt tagmondatok lineáris elrendezési lehetőségeinek elemzése. [Analysis of the possible linear arrangement of independent clauses.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 95–119.

Petőfi S. János: Szövegek linearizált szerkezetének analitikus megközelítése. [An analytical approach to the linearized structure of texts.] In: Szikszainé Nagy Irma (Szerk.) *Officina Textologica 6.* Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. (Magyar nyelvű szövegek elemzése. Diszkusszió). Debrecen, 2002. 50–68.

Petőfi S. János: Előszó. [Foreword.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 7.

Petőfi S. János: Előszó. [Foreword.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 7–8.

Petőfi S. János: A szövegorganizáció megjelenési formái és a jelentés három arca. [Different forms of text organization and the three facets of meaning.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 91–104.

Skutta Franciska: Tematikus progresszió és lineáris elrendezés. [Thematic progression and linear arrangement.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 14–33.

Skutta Franciska: Téma–réma tagolás és szintaktikai műveletek a francia összetett mondatban. [Theme-rheme organization and syntactic operations in complex sentences in French.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 51–64.

Skutta Franciska: „Gyönyörű egy okoskodás!" — A tematikus progresszió karikatúrája Molière-nél. [What a fine piece of argumentation! The caricature of thematic progression in Molière's plays.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 17–32. 33–44.

Szabó Zoltán: Linearitás és összehasonlító stilisztika. [Linearity and comparative stylistics.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 7.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: téma–réma szerkezet). Debrecen, 2002. 123–128.

Szabó Zoltán: A tematikus progresszió stílustörténeti megközelítésben. [Thematic progression in the context of the history of style.] In: Petőfi S. János–Szikszainé Nagy Irma (Szerk.) *Officina Textologica 9.* A kontrasztív szövegnyelvészet aspektusai. (Linearizáció: tematikus progresszió). Debrecen, 2003. 81–90.

Szikoráné Kovács Eszter: Egy elfogadhatónak minősíthető prózarészlet-vehikulum létrehozása. [Creation of a fairly acceptable vehiculum for a piece of prose.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 67–69

Szikszainé Nagy Irma: Tűnődés a linearitás kérdésein. Utószó. [Contemplation on the questions about linearity. Epilogue.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 132–141.

Szikszainé Nagy Irma: Előszó. [Foreword.] In: Szikszainé Nagy Irma (Szerk.) *Officina Textologica 6.* Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez. (Magyar nyelvű szövegek elemzése. Diszkusszió). Debrecen, 2002. 7–10.

Tolcsvai-Nagy Gábor: Vázlat a mondat és szórend funkcionális viszonyáról folyó magyar nyelvészeti kutatásokról. [An outline of the Hungarian

linguistic research on the functional relationship between sentence and word order.] In: Szikszainé Nagy Irma (Szerk.) *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez.* (Magyar nyelvű szövegek elemzése. Diszkusszió). Debrecen, 2002.11–24.

Vass László: Egy adott verbális szöveg formációjával kapcsolatos kreatív-produktív gyakorlat. [A creative-productive exercise in assessing the formation of a selected verbal text.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 70–77.

Wacha Imre: A prozódia szerepe a lineáris elrendeződés kialakulásában. [The role of prosody in the formation of linear arrangement.] In: *Szikszainé Nagy Irma* (Szerk.): *Officina Textologica 6. Szövegmondat-összetevők lehetséges lineáris elrendezéseinek elemzéséhez* (Magyar nyelvű szövegek elemzéséhez). Debrecen, 1999. 34–47.

5.
Connectives and discourse markers
Describing structural and pragmatical markers in the framework of textology[1]
István Csűry

This chapter is about elementary components of text/discourse to which no categorisation of universal consent has been applied yet, as it is shown by the terminological diversity in itself (*discourse particles*, *discourse markers*, *connectives*, *connectors*, etc.). Our intent is to outline the achievements of research carried out on their issues in the framework of the *Officina Textologica* (OT) project.

In the first part, publications devoted to connectives and discourse markers are reviewed. The second section discusses the main problems of identifying and classifying such elements and proposes a set of criteria that enables us to tell apart text/discourse structuring element types in a simple yet complete and useful way. Terminological issues are addressed and difficulties of empirical research are highlighted. Part three presents analyses of several text excerpts in order to demonstrate how syntactic, informational and discourse structures interact and how connectives act as an interface between them.

1. Studies related to connectives and discourse markers published in the volumes of Officina Textologica

In the original conception and the first publications of the OT project, no specific attention was paid to lexical items used as discourse markers, connectives or frame markers inasmuch as they were either considered as mere formal, logical-like means of linking sentences or simply judged to be of poor interest in a semiotic-textological approach of (written) discourse. In fact, global theoretical challenges of describing text as a complex sign as well as the large amount of work to be done on coreference and linearization hindered the issues of such, apparently isolated, phenomena. However, owing to the polyglot setting of the research program and the different scientific background of the participants, papers on structural and pragmatic markers of discourse have been published in the OT series as early as 2001.

Having treated various problems of grammar, text linguistics and textology in volume 5, the author of the present study introduces connectives in this framework as a full-fledged research topic. In (Csűry 2001b), he intends to define the category of connectives by outlining a classification of lexical and grammatical means of structuring texts. Given the lack of any comprehensive approach to connectives as well as the vagueness of their definitions in Hungarian textology, the paper mainly refers to French linguists' works, especially in pragmatics[2], manifesting a sustained attention to this category. In spite of some controversial aspects of these sources, they offer useful considerations for determining the essential features of connectives. (Csűry 2001b) points out that the latter serve to construct complex textual units as wholes by setting up semantic/pragmatic relations between explicitly and implicitly formulated contents, and form a functional rather than lexical category. Notwithstanding the cases in which they assume a connexive role as well, the primary function of connectives is to build up coherence. A series of analyses of textual examples is given at the end of the paper in order to illustrate the authors' claims, at the same time showing that phenomena related to the use of connectives are to be examined not on the sentence but on the text/discourse level.

Volume 7 is devoted to issues of linearization in correlation with information structure. In this framework, I address syntactic peculiarities of connective use on the basis of a corpus study of French and Hungarian data. These two languages differ with regard to word order: syntactic function determines the place of constituents in the former whereas in the latter, information structure decides which one is to be chosen among several possible linear arrangements. However, connectives of an adverbial nature are more or less mobile in both, and display some puzzling variations of word order. The principal claim of (Csűry 2002) is that connective position and semantic structure of text are interdependent. As it was formerly stated, connectives have an essential, semantic/pragmatic function in realizing coherence relations, and a potential one of formally linking contiguous pieces of text (or sentence) as connexity markers. While the first is effective regardless of the syntactic position of the connective, its scope varies depending on word order, and sentence- (or clause-) initial position[3] activates the second, connexity marker, function in such a way that the syntactic unit preceding it immediately is interpreted as the bearer of the meaning the given coherence relation is referred to. In other cases, the so-called left term of the connective may spread over text blocks of variable dimensions

---

[2] Adam, Anscombre, Ducrot, Moeschler, Roulet
[3] That means the leftmost possible position, constrained by the presence of other occurring terms (see section 3, page 6).

and structure, and, what is more, the immediate neighbour next to the left of the clause containing the connective may even be an embedded unit that does not directly contribute to the coherence relation it establishes. As for French, corpus research reveals massive evidence for this text and sentence building strategy, formulated by the author as the principle of the primality of connexion, and parallel phenomena were observed on a smaller sample of Hungarian.

In volume 9, which is a collection of papers on thematic progression, (Csűry 2003) discusses the question whether it is possible to draw conclusions, on the basis of the presence of certain connectors, regarding the thematic structure of their co-texts, and, reciprocally, whether the interpretation of semantic relations marked by connectives depends on the thematic structure of the surrounding text block. The answer is to some extent positive insofar as thematic shifts delimit text chunks taken in consideration as bearers of meanings linked together by the connective. Unfortunately, no clean-cut rule seems to apply to mechanically segmenting text from this point of view because of the recursive nature of thematic structures, interwoven thematic networks and the absence of mutually univoquous correspondence between (types of) textual and thematic units. The paper presents analyses of eight excerpts from texts of several types in order to demonstrate these assumptions.

Volume 10 discusses the role of conceptual schemas in constructing text. My contribution consists in examining the use of explanatory and combinatorial lexicology[4] and of hypertextual linking of lexical representations for the sake of investigating conceptual schemas at work in building text meaning. In this perspective, connectives confront lexicologists with special issues but, in a textological approach, their procedural meaning can be successfully analyzed in light of such representations of meanings in context. Reversely, by virtue of their procedural meaning, connectives make emerge latent conceptual schemas in context, as it is showed by (Csűry 2004).

Volume 13 is entirely devoted to connectives: this monograph, intituled *Kis könyv a konnektorokról* (*Small Book on Connectives*) summarizes the results of research carried out till then, marks orientations for future investigations and demonstrates the indispensability of a corpus linguistic approach. (Csűry 2005) is intended to be a contribution to working out definitions and problems concerning connectives by placing all linguistic issues to be raised in the adequate, i. e. textological, framework of investigation.

Chapter 1 deals with the notion of connectives and with the peculiarities of their function and use. After a survey of several approaches, connectives are defined (partly according to relevance theory) as a function in the text structure fulfilled by units with a procedural meaning and used for optimizing information

---

[4] For a detailed presentation of the theory, see (Mel'cuk, Clas, and Polguère 1995).

processing. This category is identified as one of the subsets of complete text constituent marking functors and to be considered in the largest category of text structuring elements. I also point out the anaphoric nature of connectives and argue for a multi-level representation of the relation of conjuncts.

In chapter 2, I examine the place of connectives in the linguistic system. A certain ambiguity was unavoidable at this point inasmuch as I intend to characterise not only the function of connectives themselves but also the classes of linguistic items usually appearing in this function. In this sense, I have to deal with the semantics of connectives, i. e. the problem of relation types and the interpretation instructions conveyed by different lexical items. As an example, one of the basic semantic relations marked by connectives, namely, contrast, is defined and described at some length, in order to present its underlying relationships as well as its surface realisations. Contrast can be defined in an exact way as a binary relation of concepts founded on negation. The way contrast-marking connectives mobilize underlying contents (implications or expectations) is also demonstrated. Since it is often difficult to access these explicitly non-manifest contents, there are several approaches to the interpretation of contrast. Apart from criticizing the argumentativist and the syllogistic ones, I emphasize the role of conceptual schemata.

In chapter 3, I focus on text structure. After an analysis and demonstration of how partial semantic structures of discourse may be jumbled, I take into account the difficulties and possibilities of localizing and delimiting the so-called poles of connectives, i. e. the portions of context which bear the meanings they link together. I describe an XML-based annotation schema the use of which permits to put further research on connectives on an empirical basis by the use of appropriately tagged corpora. I look for semantic and formal criteria in the text structure which are necessary to identify discourse constituents connected by connectives in a consistent and unambiguous way. I also reconsider the lexical characteristics of items in the role of connectives in order to highlight their specific potential for the articulation of a given context.

Volume 14 resumes the discussion on conceptual schemas and focuses on the role of scripts or scenarios in constructing and interpreting texts. In (Csűry 2005), I examine the place and function of connectives in dialogues, emphasizing the double way they may link semantic contents, the one being anaphoric (co-textual) and the other deictic (contextual). This possibility allows connectives not only to bridge units of meaning inside a single turn or belonging to different speakers' turns but to establish coherence relations as well with any element of context, including cognitive elements of the speakers' (supposed) common ground. Thus, reference can be made to scripts that are likely to underlie the current interaction. The paper starts with a review of types of text structuring elements that may be used as cues in dialogue

production/interpretation. In parallel, dialogue phases are distinguished and some basic distinction is made of dialogue types as well in terms of the formal or institutionalized character of their scenarios. On these grounds, (Csűry 2005) claims that dialogic use of deictic connectives as indirect means of structuring scenarios is common mostly in internal sequences of informal dialogues. Their functions consist, on the one hand, in ensuring that the realization of the scenario carry on normally and, on the other hand, in sanctioning and/or correcting any deviation from its expected fulfillment. As for the frequency of this type of connective use, it seems to be rather limited with respect to the overall number of words in dialogue texts, especially in comparison with explicit dialogue structuring means in dialogues that follow formal scenarios. By the same token, (Csűry 2005) suggests that the term of script/scenario should refer to two kinds of analogous structures belonging to different levels: a given type of a communicational event has a scenario as a whole, composed of a limited number of more general partial scripts, such as questioning and answering, or treating a misunderstanding. Deictic connectives may signal scripts/scenarios of both levels.

From the point of view of our concern, Volume 16 is, after Volume 5, another landmark in the evolution of the Officina Textologica project. In fact, two papers of the seven published in this volume address issues of classification and treatment of text structuring items as well as of their relationship with different aspects of cohesion and coherence. (Furkó 2011) leads the reader on the slippery ground of the so-called discourse markers by providing an evaluative overview of their study in the relevant English literature. He concludes that terminology is not unified and there is no generally accepted typology, concurring with (Csűry 2001b), who came to the same conclusion with respect to connectives. Furkó looks at the functions discourse markers have on different planes of discourse as well as the role they play in connectivity. He points out a range of uncertainties and unclarified issues the resolution of which is crucial with a view to formulating a unified approach to cohesion and coherence in general and discourse markers in particular.

The other study in question, (Csűry 2011), recapitulates the main advances of the *Officina Textologica* project in the field of connectives and the principal concerns they present with respect to semiotic textology and research on discourse markers. As for the latter, it is suggested in the introductory part of the paper that the term should be considered as a denomination of the broadest category of items structuring (meaning in) discourse, which allows us to see connectives as a subcategory in this framework, facilitating further discussion. The author then turns to French linguistics/pragmatics as an abundant source of knowledge on connectives that present several clean-cut orientations and may provide Hungarian research in textology with useful points of reference. After

defining nine criteria according to which the different approaches as well as the results they have permitted to obtain can be compared, he gives an overview of argumentation theory (Anscombre, Ducrot and others), conversation analysis (Eddy Roulet and his Geneva School), relevance theory (Moeschler, Reboul and others) and some other grammatical, semantic, text linguistic, corpus-based and diachronic approaches (Adam, Morel, Lamiroy and Charolles, Combettes). Finally, he outlines the major claims of the relevant Hungarian literature and points out the lack of a generally accepted theoretical synthesis founded on solid empirical evidence, and suitable for language description tasks. The paper concludes with programmatic statements concerning the discourse marker/connective branch of the *Officina Textologica* project in view of the fact that the multidisciplinary character of semiotic textology as a theoretical framework seems to fit such research. (Csűry 2011) urges that a comprehensive theoretical revisiting of the field be carried out, combined with empirical investigations, and that research should be extended to oral discourse.

2. Some basic problems of the study of text/discourse structuring element types

As a starting remark, we should stress that only the special category of connectives has made an object of research followed in the OT project, and not all possible kinds of text/discourse structuring elements, although such an extension is, undoubtedly, inevitable, given the global aims of textology. In what follows, let us therefore concentrate principally on connectives, making , at the same time, necessary remarks on the neighboring categories, especially as the latter need to be taken into account for a proper definition of the former.

Discourse is organized in hierarchically ordered semantic constructions of variable dimensions and complexity, composed of recursive elementary structures, and textured with lexical and grammatical relations, such as co-reference or tense relations, extending beyond sentence borders. Furthermore, these constructions are usually, but not always, articulated by different kinds of lexical items or even by complex expressions specifically used for indicating the makeup of their meaning. The relevant literature proves a high interest of researchers in this field.[5]

Let us consider first of all the linguistic items to which the most permissive conception of connecting means of text/discourse might apply. It should be indeed reasonable to identify them as a whole set of elements ensuring

---

[5] See, for example, (Charolles 1997), or the thematic issues of the journal *Discours*: *Approches fonctionnelles de la structuration des textes* (Ho-Dac and Bolly 2011), *Multidisciplinary Perspectives on Signalling Text Organisation* (Ho-Dac et al. 2012).

connexity, without reference to the status and/or position of the members of the relation they mark. One could designate this largest category, marked *1* in the figure below, as ***text structuring elements***; it embraces, among others, grammatical forms denoting possession or temporal relations, e.g. In other words, every linguistic item used expressly in order to *ensure connexity* and/or as marks of semantic and/or pragmatic relations in text/discourse on the level of the *significans* (more precisely, in the *notatio*[6]) is a text structuring element.
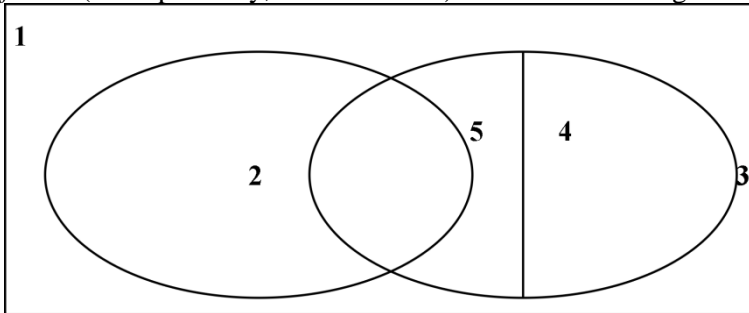

Figure 1

One of the subsets in this category, marked *2*, is constituted of ***connexive elements*** that are *used to ensure linear continuity* of text by means of marking structural and semantic relations of contiguous simple macro-architectonical units. Such a role may be played by a conjunction, a pronoun or an adverb.

The other subset, marked *3*, is that of ***text organizers marking/creating complex textual units as wholes***. Words and syntagms[7] in this function have for common feature a key role in *articulating* the *significans* of the text *according to the semantic relationships* the text has to express.

This subset has to be divided in turn so as to differentiate two further subsets. The first, marked *4* in the figure, contains items ***marking linearly composed arrays of text units***. These mark either the place or the (semantic/pragmatic) role of a macro-architectonical text unit (of the order of sentence, sentence group or sequence) in the overall structure of a given text. Their effect therefore takes place at the level of the representation of the referent (or rather that of *relatum-imago*). Lexical items used in this function have a proper lexical meaning providing them some degree of referential autonomy. The following belong to this category:

- ***linear integration markers*** (e. g. *on the one hand … on the other hand*; *the first … the second*, etc.);

---

[6]   Terms of the Petőfi model of text (passim), especially (Petőfi 1996: 12), (Petőfi and Benkes 1998: 41) and (Petőfi 2004: 27 sqq.).

[7]   Let us notice as a morphological property of this category that it contains no bound morphemes.

- ***repetition and reformulation markers*** (e. g. *in other words*; *in sum*, etc.);
- ***markers and predicates of universe of discourse*** (examples of the former: in this year, …; according to Steve, …, etc.; examples of the latter: *once upon a time…*; *suppose that…*, etc.).

The second subset, marked *5* in the figure, is that of **connectives**[8]. A connective signals that the text constituents it connects stand in a specific semantic (semantic-logic, semantic-pragmatic) relationship by activating a certain interpretational-inferential procedure. It follows from this property that, in most cases, connectives relate physically present units of text to explicitly not manifested components of discourse meaning that text needs to be completed with in an explicit form in order to provide exact semantic analysis with a suitable object. The function of connective is held by verbal entities without referential autonomy (conjunctions, certain adverbs, some complex lexical units and syntagms) the presence of which allows for an unambiguous interpretation of the linked text constituents (and, therefore, that of the text).

We need to insist on the fact that the categories having been enumerated are not lexical or syntactical but **text structuring functions**, even if a set of lexical-grammatical units or another has a privileged role in assuming them. That is to say, in conceiving connectives as a collection of lexical items, we must keep in mind that we are dealing with an open set containing some clearly defined core elements as well as (more vague) peripheral ones. The latter are polysemic because they get, in some contexts, a new, procedural meaning while keeping elsewhere their original, referential one. An adverb, for example, generally used as a verbal adjunct, may be uttered as a connective being promoted to the function of a sentence adverb while no specific contextual (syntactic) feature indicates the semantic difference. Let us illustrate this issue with French adverb *maintenant* 'now'.

(1) Luc est devenu riche. Maintenant, il peut s'acheter un yacht. *(Luc became rich. Now, he can afford a yacht. — temporal relation)*

(2) Luc est devenu riche. Maintenant, est-ce qu'il faut en déduire qu'il est malhonnête? (*Luc became rich. Now, may we conclude that he is immoral? — sentence adverb in connective function; concession.*)

---

[8] We might call *connectives* the whole set of all text structuring elements and introduce another term to designate this particular subset. This solution would be more or less consistent with the abovementioned definitions. However, it is especially about this subset and its elements that one can find abundant literature, where the term *connective* generally occurs in this narrower sense.

At the same time, a term should never be called a marker, an organizer or a connective in general, only its utterances considered in (con)text. Furthermore, it should be pointed out that these categories may have a common subset. For instance, an item occurring with a connective function (i. e. having a semantic/pragmatic role) may, in the same time, ensure connexity, as in (3). Obviously, as it is illustrated by (4), the two functions do not necessarily co-occur (in this example, connexity is marked by the underlined part of the sentence).

   (3) For many years we have been building a program to give the farmer a reasonable measure of protection against the special hazards to which he is exposed. That program was improved at the last session of the Congress.
HOWEVER, our farm legislation is still not adequate. *(Compleat Lexical Tutor v.6.2, corpus "Presidential speeches")*

   (4) The first information in an authentic form from the agent of the United States, appointed under the Administration of my predecessor, was received at the State Department on the 9th of November last.
This is contained in a letter, dated the 17th of October, addressed by him to one of our citizens then in Mexico with a view of having it communicated to that Department. From this it appears that the agent on the 20th of September, 1844, gave a receipt to the treasury of Mexico for the amount of the April and July installments of the indemnity.
In the same communication, **HOWEVER**, he asserts that he had not received a single dollar in cash, but that he holds such securities as warranted him at the time in giving the receipt, and entertains no doubt but that he will eventually obtain the money. *(Compleat Lexical Tutor v.6.2, corpus "Presidential speeches")*

One should have also noticed that Figure 1 is somewhat misleading in the light of the last remarks as it suggests that there are preexistent lexical sets of text structuring elements as such. The relationship of these functions might be better represented with graphs, like in Figure 2.
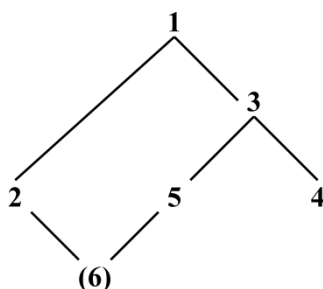
Figure 2

Let us now take a closer look at terminological issues as the very first problem related to text structuring elements is a terminological one. As for the semantic/pragmatic structures of discourse, Charolles (1997) introduces a fairly comprehensive and widely accepted taxonomy and terminology; however, given its use in cognitive semantics, the key term of *cadre* 'frame' might lead to ambiguities in a multilingual context. Consequently, these structures are rather called *discourse frames* in English. Derivates like *framing adverbials* (and their equivalents in other languages) are less ambiguous and, therefore, could be generally used referring to a class of structural markers. These correspond to our class 4, i. e. the markers of linearly composed text arrays, and establish ***forward-looking relations***. The question of how to refer to the whole category of text structuring elements remains unanswered, elements, which are instantiated, on the one hand, by conjunctions and other, more or less complex, expressions and, on the other hand, by items that do not or not always mark units of meaning one should call "frames". In particular, this is the case of marks/markers of ***backward-looking relations***, such as connectives, the anaphoric nature of which is apparent. The equivalents of the term *connective* are used in several languages (e. g. *connecteur* in French, *Konnektor* in German and in Hungarian) with various implications: it may denote a logical, a grammatical, a textological or a pragmatical approach or some specific combination of these factors. This term is often a part of terminologies that cover a differentiated set of means of structuring or marking text/discourse on the verbal, semantic and pragmatic levels. In this way, *discourse marker* and its translations may be used with reference to all kinds of items adapted to such functions. However, this term seems vague to some extent due to its controversial views in pragmatics (for instance, (Schiffrin 1988; Fraser 1990; Redeker 2006)). Nevertheless, even if considering these terminological problems unresolved, one might refer to the widest category of linguistic means playing a role in (de)marking structural units of text/discourse on the semantic/pragmatic level as *discourse markers* (DMs) and reserve the term *connective* (C) for one of its possible subcategories, distinguished from the others by (at least) their procedural meaning in logical-

like constructions of sense. In fact, it is the very vagueness of the term *discourse marker* that allows such an extended usage, whereas the above mentioned term *text organizers marking/creating complex textual units as wholes* is, despite its descriptive exactness, rather lengthy and inconvenient for current use. At the same time, it is quite clear that confusions might result from this terminological choice as well since expressions that have barely anything to do with our functional categories are often called DMs.

Evidently, there are underlying theoretical problems in addition to this terminological diversity. As a matter of fact, structures of text/discourse are described in various theoretical frameworks that results in heterogeneity of perception, categorization and terminology of the linguistic items involved. Moreover, different research projects focusing on similar or the same subjects often do not make reference to each other, especially if researchers belong to different language areas (English vs. French e. g.). Such discrepancies are also due to the fact that the lexical items under scrutiny do not constitute a homogenous class, their uses and effects are varied and, what is more, they do not necessarily and invariably mark a particular discourse relation. In the relevant literature, either the category of DMs is treated from a fully theoretical point of view, without any substantial empirical support, or empirical analyses follow a restrictively predetermined orientation that excludes in advance the taking into account of phenomena that do not fit the given theoretical framework. In most cases, the (sets of) items under scrutiny are considered separately from the others, and even if attested examples of language use replace simple introspection, their sources are restricted to particular types of discourse. For instance, the description of French *mais* by Luscher (1994) illustrates the application of relevance theory, papers written in the framework of successive versions of Anscombre's and Ducrot's argumentation theory examine individual Cs, [9] Hungarian analyses of text meaning structures (like Békési (1993)) are based on well-styled written discourse while works following the Anglo-American DM tradition focus on spoken interaction. All these are valuable contributions to a better knowledge of how semantic-pragmatic text/discourse relations can be marked, identified and interpreted but, as a consequence of research-methodological choices, the descriptions are, for the most part, partial, incommensurable, suffer from empirical weaknesses or are simply lacking. In our own work[10], we have pointed out that previous research along these lines has yielded partial results; further progress can only be made by following an integrative approach.

---

[9] For instance, (Anscombre 1983; Anscombre and Ducrot 1983).
[10] (Csűry 2001a; Csűry 2005) among others.

Beside the general theoretical problems, several particular issues arise as well. First of all, we need data in order to confirm hypotheses and ground analyses in DM research. As it was highlighted by Péry-Woodley (2005: 185), "Studies on discourse are actually characterized by a qualitative approach of small amounts of data by means of manual, thus subjective, methods […] which creates an obstacle […] to the generalization of their results. […] We are in need of techniques allowing us to apprehend and to articulate often confounded mechanisms at different levels of granularity"[11]. In recent years, indeed, one can observe a growing interest in discourse-level corpus linguistics. However, several problems arise from the very first steps of corpus analysis in this particular domain inasmuch as structures to be observed do not fit a unique and clearly describable pattern. Furthermore, computerized processing of linguistic data is a very complex issue since it is not just *words* we have to look for in corpora but *meaningful units* of variable dimension and structure, which are not only contiguous but may display embedded and overlapping arrangements, and *their relations* in a functional perspective. Thus, formal clues for computerized processing of such textual/discursive data form a matter of research in themselves. What is more, taking account of live communication involves similar efforts in the field of processing multimodal signals.

The other particular issue consists in the systematic description of lexical/phraseological items habitually occurring as DMs. Obviously, dictionaries and descriptive grammars give basic (or, in some cases, even somewhat more sophisticated) information about them and we have case studies as well as analyses of particular sets of items at our disposal. Despite this, given the divergences of orientations and the partial character of research described above, there is a lack of synthesis on DMs considered in their mutual relationship and with respect to every relevant aspect of their use, founded on solid empirical investigations. Not surprisingly, contrastive analyses are far from being systematic in spite of delicate problems of equivalence related to DMs; a comprehensive study like (Rudolph 1996) is a rare exception.

Our basic assumption is that the study of DMs is of the utmost importance as far as they guide the inferential process of interpretation of text/discourse. Furthermore, we are convinced that inquiry into linguistic phenomena related to meaning and interpretation necessitates a global approach, i. e. the study of all kinds of language use (written/spoken, monologic/dialogic, etc.). Our third contention is that the markers of relations in semantic-communicative structure,

---

[11] Our own translation from French.

whatever their nature might be, can only be studied effectively in terms of their interrelations and in context. As previous research suggests, we assume that connectives and other discourse markers do not constitute a lexical class but form a functional category. Finally, we think that some open sets of lexical items, with quasi-paradigmatic inner relationships, could be identified and described on the basis of their regular use as DMs. A corpus-based contextual approach taking into account large pieces of discourse seems to be the most appropriate way to establish a complete and coherent description of these lexical items as well as of discourse relations marked by them.

3. Connectives at the interface of syntactic, informational and discourse structures

The study of connectives is an interdisciplinary one as it is situated in the intersection of lexicology, syntax, text/discourse analysis, semantics and pragmatics. Previously, summarizing the principal claims of (Csűry 2002) and (Csűry 2003), we saw that the position of a connective in the sentence structure is closely related to the thematic structure of its context, and, consequently, influences the way in which the actual discourse sequence will be interpreted. Difficulties arising from embedded relations and recursive structures have been signaled as well. In this section, let us briefly expose, with examples at hand and referring to analyses made in the framework of the OT project, how grammar and meaning interact in the context of connectives and how an explanation of apparently unmotivated word order variations has emerged from textological research.

A fundamental problem the analyst is confronted with is the difference between syntactic and semantic/pragmatic structures. Many difficulties arise from the unpredictable character of the linguistic structures that connectives can connect. While some items are characterized by contextual restrictions, most of them admit as terms[12] words or word groups, phrases, clauses, sentences, and multi-sentence discourse sequences (i. e. macrostructural units) as well. Moreover, the structural status of the connected terms may be different and, of course, nothing allows to preview the amount of discourse making a term. Generally speaking, the discordance of syntactic and semantic relations makes

---

[12] According to (Csűry 2001a), three levels of analysis should be distinguished for the sake of a proper interpretation of coherence relations established by connectives. While they appear at articulation points of *verbal blocks*, semantic entities (their *poles*) to be taken into account for the interpretation of the given coherence relation are frequently carried (explicitly or inferably) by only parts of these verbal blocks. These parts (of texts or sentences) should be called the *terms* of the connective.

rather the rule than the exception on discourse level. Finally, the linear arrangement of discourse constituents does not indicate what belongs to a certain coherence relation and what constitutes an external element in a given context. Several connectives admit distant terms, that is to say, the clause, sentence or sequence on the left side of the connective are not obligatorily its first term since parenthetical discourse constituents may occur at this point. Furthermore, we need to count with overlapping relations as well since text structure is made up of a complex set of imbricate relations or frameworks of different kinds. Syntactic mobility of sentence adverbials acting as connectives as well as semantic effects of their utterance are to be accounted for with respect to all these factors.

In French, both positions of *par contre* and *cependant* (representing here numbers of adverbial connectives that have a similar behavior) shown in examples (5)-(8)[13] are acceptable and no clearly conceivable difference seems to result of this variation as long as analysis does not exceed the boundaries of the sentence.

(5) **PAR CONTRE**, Duroy dînait tous les jeudis dans le ménage et faisait la cour au mari en lui parlant agriculture.

(6) Duroy, **PAR CONTRE**, dînait tous les jeudis dans le ménage et faisait la cour au mari en lui parlant agriculture.

(7) Le nouveau gouvernement a **CEPENDANT** entrepris un programme de grande ampleur.

(8) **CEPENDANT**, le nouveau gouvernement a entrepris un programme de grande ampleur.

However, once considered in a given context, the different word orderings are not equally likely to occur since they prove not only to affect meaning but also to display , at times,differences of acceptability: authors (and speakers) probably have some specific reason to follow a particular pattern when placing connectives in sentence structure. Here is the source of (6) (and of (5) which is derived from the former), an excerpt from a novel of Maupassant:

(9) Il habitait maintenant rue de Constantinople, où il avait transporté sa malle, sa brosse, son rasoir et son savon, ce qui constituait son déménagement. Deux ou trois fois par semaine, la jeune femme arrivait avant qu'il fût levé,

---

[13] French examples, used for having abundant data at our disposal obtained by extensive corpus research, are not translated on purpose at this point.

se déshabillait en une minute et se glissait dans le lit, toute frémissante du froid du dehors.

Duroy, **PAR CONTRE**, dînait tous les jeudis dans le ménage et faisait la cour au mari en lui parlant agriculture ; et comme il aimait lui-même les choses de la terre, ils s'intéressaient parfois tellement tous les deux à la causerie qu'ils oubliaient tout à fait leur femme sommeillant sur le canapé.

Putting *par contre* in sentence initial position would disorient the reader at the moment as he would expect some complementary or contrasting information with regard to that what was given so far about the usual behavior of the young woman when arriving to Duroy's, while this is not the case: it is oddly contrasting peculiarities of their reciprocal visits that we feel emphasized. In English, it is rather difficult to find a perfect equivalent of *par contre*: in some cases, *in contrast* gives a satisfactory solution, in other cases, *on the other hand* is a functional equivalent, but ad hoc translations can only render its actual value in the French source. Unfortunately, the translator of an English version widely spread on internet has considerably abridged the episode by omitting its ironic and erotic elements; however, s/he has maintained the original semantic structure. Reciprocal visits are put in (a weak) contrast while, quite surprisingly, we find a somewhat abusive explicitation of the value of the connective *par contre* to which a final clause corresponds in the translated text:

(10) Duroy moved his effects to the apartments in Rue de Constantinople. Two or three times a week, Mme. de-Marelle paid him visits. Duroy, ***to counterbalance them***, dined at her house every Thursday, and delighted her husband by talking agriculture to him.

(Literally: *Duroy was now living in the apartments in Rue de Constantinople where he had transported his trunk, his brush, his razor and his soap, that was what constituted his moving house. Two or three times a week, Mme. de-Marelle arrived before he would get up, she undressed herself in a minute and slipped in the bed, shivering of outside cold. As for Duroy, he dined at her house every Thursday,* etc.)

Similarly, we understand at first glance why only (7) can correctly depict the actual state of affairs when we replace it in its original context:

(11) La Roumanie doit affronter un défi particulier en matière d'affaires intérieures et de justice. Elle a pour l'instant fait des progrès limités en matière de reprise de l'acquis en ce domaine. Le nouveau gouvernement a **CEPENDANT** entrepris un programme de grande ampleur pour mener à bien les réformes institutionnelles indispensables.

With (8) in the same context, one would suggest that limited progress has been made in spite of governmental efforts. In reality, it is former progress that is being qualified unsatisfactorily while the new government's program is being opposed to this situation. This is the way we can interpret the English version as well[14]:

(12) Romania faces a particular challenge in justice and home affairs. So far it has made limited progress in taking on the *acquis* in this field. The new government has <*AT THE SAME TIME*> undertaken an ambitious programme *(sic!)* to introduce the essential institutional reforms.

Let us now consider the aforementioned structural issues in the light of which semantic effects of connective position could be better understood. The following English examples not only facilitate the task of demonstration but also show the main cross-linguistic, if not universal, characteristics of connectives from the point of view of syntax and semantics.

Usually, we represent the use of connectives with examples in which they mark the relation of two clauses or sentences. It is indeed one of the basic configurations of structures of meaning in discourse, as shown by (13) (the numbers in brackets serve to identify syntactic units that are semantic blocks of a concessive realization of contrast at the same time):

(13) [1]Although Sam Rayburn affects a gruff exterior in many instances, [2]**NEVERTHELESS** he is fundamentally a man of warm heart and gentle disposition.

---

[14] Typography marks that the official English text contains no connective at this point. Given that the documents of the European Union are published in all official languages of the Union, we cannot establish if the French or the English version (or a third one) is the original. Either the difficulty of reproducing the exact value of a connective in an other language has lead the translator of a French source text to choose not to use any connective at all in the English version, or the intention of clearly expliciting textual relations in the French translation of an English original has made the translator introduce a connective. The spelling mistake allows for supposing at the same time that the English version was not written by a native speaker. According to our experience in language teaching, learners of foreign languages acquire a differentiated use of connectives only at a very advanced stage and often hesitate even as fluent users.

For the sake of comparison, it should be useful to give a formal, graphical representation of this construction (Figure 3).[15]
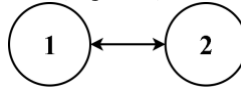

Figure 3

However, as relations marked by connectives may extend beyond sentence boundaries, we have to be conscious of the fact that a sentence containing such an element has a particular position in the meaningful structure of discourse and this must be taken into account even if the configuration is similar to the basic one seen above. In (14), we find the same connective (*nevertheless*) in unit 4 in sentence initial position:

> (14) [1]Presumably, if the reverse is the case and the good effect is more certain than the evil result that may be forthcoming, not only must the good and the evil be prudentially weighed and found proportionate, but also calculation of the probabilities and of the degree of certainty or uncertainty in the good or evil effect must be taken into account. [2]There must not only be greater good than evil objectively in view, but also greater probability of actually doing more good than harm. [3]If an evil which is certain and extensive and immediate may rarely be compensated for by a problematic, speculative, future good, by the same token not every present, certain, and immediate good (or lesser evil) that may have to be done will be outweighed by a problematic, speculative, and future evil. [4]**NEVERTHELESS**, according to the traditional theory, a man begins in the midst of action and he analyzes its nature and immediate consequences before or while putting it forth and causing these consequences. [5]He does not expect to be able to trammel up all the future consequences of his action. [6]Above all, he does not debate mere contingencies, and therefore, if these are possibly dreadful, find himself forced into inaction.

As it is shown in Figure 4, the sentence in question is in a central position in the semantic-functional structure of this text fragment. Although we can interpret the chain of [3] and [4] separately from the rest, the former is tightly connected to [2] and [1] whereas the latter is developed by [5] and [6].

---

[15] For lack of space and for the sake of simplicity, we do not specify in the following analyses the exact nature of semantic-functional relationship of all blocks of meaning. These RST-like relations, marked by arrows and horizontal/vertical arrangement, should be read intuitively.

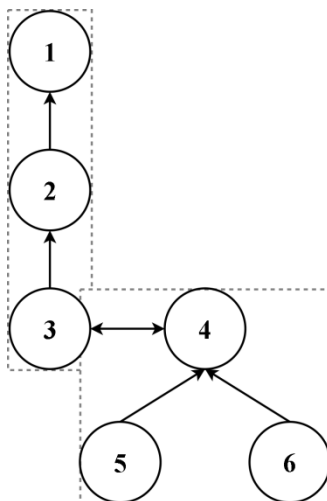*Nevertheless* marks a concessive relation between [3] and [4] and, indirectly, between the two blocks.

Figure 4

Conversely, connectives are often used in order to mark coherence relations between contents linked to the constituents of the same sentence. In such cases, semantic structure established by the connective does not necessarily map the syntactic one. From the point of view of constructions of meaning of discourse, context should not be neglected, either. In (15), *nevertheless* marks a concessive relation between two adjectives qualifying the same referent (*example*), i. e. puts the attribution of the qualities they denote on the same level of structure of meaning while they occupy quite distant points of syntactic structure, as shown in Figure 5.

(15) This understanding provides a very simple example of the fact that one can eliminate fear without instituting any controls. In fact, although we have dispelled the fear, we have not necessarily assured ourselves that there are no dangers. There is still the remote possibility of planetoid collision. A meteor could fall on San Francisco. Solar activities could presumably bring long periods of flood or drought. Our understanding of the solar system has taught us to replace our former elaborate rituals with the appropriate action which, in this case, amounts to doing nothing. Yet we no longer feel uneasy. This almost trivial example is **NEVERTHELESS** suggestive, for there are some elements in common between the antique fear that the days would get shorter and shorter and our present fear of war.
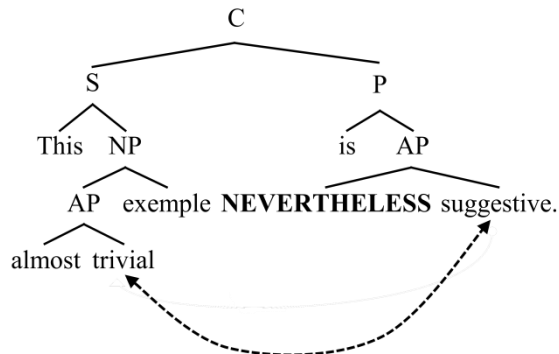
Figure 5

The role of context and of overall semantic structure of discourse can be clearly observed in example (16) which illustrates as well that sentences are not to be automatically considered as basic units of the semantic/pragmatic structure of discourse. It is once more the same connective that marks a concessive relation in this fragment but the sentence in which it occurs, [5], stands in contrast with a clause at some distance backwards, denying or at least restricting expectations that one might draw from it as conclusion due to its argumentative power. This relation is quite obvious given their lexical relatedness (the verb *communicate* of the first unit being replicated in the nominal form *communication* in the second). But (as we can see in Figure 6) there are two other units, [3] and [4], inserted between [2] and [5]. In the former the meaning of [2] is further developed whereas in [4] the author ends quoting directly his source and adds to the block formed by [2] and [3] some information on a previous state of affairs. In other words, the coherence relation marked by *nevertheless* passes over the first two units to the left of [5] implicated in this relationship only in an indirect way.

(16) [1]Hildreth states that, "[2]As an interactive system the online catalog can dynamically communicate with its user, [3]it can be responsive and informative at a given time to a given need" [4]all of which was not possible in previous catalogues. [5]In present systems the level of communication is, **NEVERTHELESS**, limited and superficial.
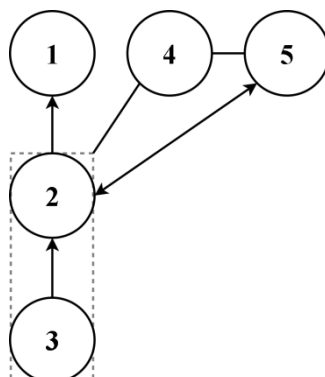
Figure 6

It is precisely in such intricate constructions of meaning, containing embedded units between the poles of a relation marked by a connective, that one finds adverbial connectives in non-initial position. In order to avoid misleading local connections between neighboring text units, writers (and probably speakers) put a thematic marker or a framing adverbial in sentence initial position and relegate the connective to specific adverbial positions inside or at the end of the sentence. The last two examples of the series (again the same text in two languages) demonstrate this phenomenon with a rather complex text structure represented in Figure 7, containing coordinate and subordinate units as well and three whole sentences embedded between the poles of the relation marked by the connective. This time we find a connective in both versions and, what is more, in the same position. In (17) and (18), *toutefois* and *however* mark the denial of the expectation that prices are the main factor of competitiveness, suggested by [3].

(17) [1]Pour permettre à l'agriculture européenne de profiter de l'évolution a priori positive du marché mondial, [2]une nouvelle réforme de la PAC doit améliorer la compétitivité de l'agriculture européenne sur les marchés tant intérieurs qu'extérieurs. [3]L'abaissement des prix profitera aux consommateurs et laissera une plus grande marge pour une différenciation des prix en faveur des produits de qualité supérieure. [4]Une orientation accrue des activités en fonction des impératifs du marché facilitera l'intégration progressive des nouveaux États membres [5]et contribuera à la préparation de l'Union aux prochaines négociations dans le cadre de l'OMC. [6]Elle aidera aussi l'Union à consolider sa position de grande puissance exportatrice mondiale.

[7]Les prix ne représentent TOUTEFOIS qu'un aspect de la compétitivité.

(18) [1]In order to help European agriculture take advantage of the expected positive world market developments, [2]further reform of the CAP must improve the competitiveness of Union agriculture on both domestic and external markets. [3]Lower prices will benefit consumers and leave more room for price differentiation in favour of high quality speciality products). [4]Greater market orientation will facilitate the progressive integration of new Member States [5]and will help prepare the Union for the next WTO Round. [6]It will also help the Union to reinforce its position as a major world exporter.

[7]Prices are, HOWEVER, only one aspect of competitiveness.
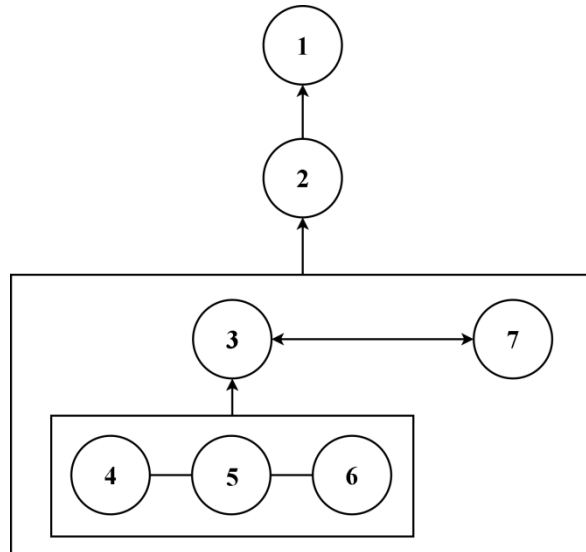
Figure 7

4. Perspectives and directions for future research

Although connectives are small elements of discourse, their study is of the greatest importance owing to their key role in text meaning and coherence relations. However, there is still plenty of work to do in their research. First, as it was mentioned above, there is no universally acknowledged definition of what should be meant by *connective,* and one can find mostly partial approaches whenever phenomena related to this category become the subject of any research. Findings and descriptions of the function of connectives and lexical elements in this role contain many inaccuracies and unfounded generalizations.

Future research on connectives and discourse markers in the framework of the OT project has three mutually dependent aims. The first aim is to answer some general theoretical questions with regard to linguistic items marking the semantic-pragmatic structures of a variety of texts/types of discourse. Critical syntheses and discussion of theoretical approaches to DMs and relevant literature that were carried out so far are an obvious starting point of the research in this field. There is still a lot of work to do in order to elaborate generally admitted principles of the systematization, typology, uniform categorization and the corresponding Hungarian terminology of the linguistic items under scrutiny, which might serve as a common frame of reference for research not only in textology but also in the fields of linguistics and pragmatics in general.

In addition to the integration and synthesis of theoretical principles, our research is also aimed at solving corpus linguistic problems pertaining to the empirical basis of the study of the functional category under scrutiny. Since it is imperative that our theoretical framework should be substantiated by empirical data and that we provide the necessary sources and tools for further (especially descriptive) research based on the same principles, we also have to aim at finding solutions to the theoretical and practical problems of corpus-based connective and discourse marker research, with special reference to the development of corpus analysis tools and methodologies. According to the results of our research obtained so far, we expect that lexical features of items used as DMs as well as contextual clues will prove useful for (semi-) automatic segmentation of textual units of meaning in pre-processed (POS-tagged, syntactically annotated) corpora. In processing spoken discourse, intonation and non-linguistic modalities are likely to have a similar role that has also to be studied.

References

Anscombre, Jean-Claude. 1983. „Pour autant, pourtant (et comment): à petites causes, grands effets." *Cahiers de linguistique française* (5): 37-84.

Anscombre, Jean-Claude and Oswald Ducrot. 1983. *L'argumentation dans la langue*. Liège - Bruxelles: Pierre Mardaga Éditeur.

Békési, Imre. 1993. *Jelentésszerkezetek interpretációs megközelítése*. Szeged: JGYTF Kiadó.

Csűry, István. 2001a. *Le champ lexical de* mais *: étude lexiko-grammaticale des termes d'opposition du français contemporain dans un cadre textologique*. Studia Romanica de Debrecen, Series Linguistica. Vol. 7. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2001b. „Hogyan osztályozzuk a szöveggrammatika kötőelemeit? (A konnektorok mibenlétéről és hovatartozásáról)." In *Grammatika -*

szövegnyelvészet - szövegtan [Grammar - text linguistics - textology], edited by János S. Petőfi and Irma Szikszainé Nagy, 27-33. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2004. „A kérdés: (hogyan) készítsünk furcsa szótárakat?" Chap. 8., In *A szövegorganizáció elemzésének aspektusai. (Fogalmi sémák)* [Aspects of analysis of text organization. Conceptual schemata], edited by János S. Petőfi and Irma Szikszainé Nagy, 93-106. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2005. *Kis könyv a konnektorokról*. Officina Textologica Vol. 13. [Small book on connectives]. Debrecen: Debreceni Egyetem, Magyar Nyelvtudományi Tanszék.

Csűry, István. 2011. „A konnektor mint diskurzusjelölő-osztály tárgyalásmódjai." In *A szövegösszefüggés elméleti és gyakorlati megközelítési módjai : poliglott terminológiai és fogalmi áttekintés: Petőfi S. János 80. születésnapjára* [Theoretical and empirical approaches to coherence. A polyglott overview of terminology and concepts. In honour of the eighty-year-old János S. Petőfi], edited by Edit Dobi, 95-107. Debrecen: Debreceni Egyetemi Kiadó.

Csűry, István. 2003. „Konnektorok és tematikus progresszió." In *A kontrasztív szövegnyelvészet aspektusai : linearizáció: tematikus progresszió* [Aspects of contrastive text linguistics. Linearization: thematic progression], edited by János S. Petőfi and Irma Szikszainé Nagy, 65-80. Debrecen: Kossuth Egyetemi Kiadó.

Csűry, István. 2002. „A tematikus struktúra és a konnektorok szintaxisa a franciában és a magyarban." In *A kontrasztív szövegnyelvészet aspektusai: linearizáció: téma-réma szerkezet* [Aspects of contrastive text linguistics. Linearization: theme/rheme structure], edited by János S. Petőfi and Irma Szikszainé Nagy, 75-92. Debrecen: Kossuth Egyetemi Kiadó.

Fraser, Bruce. 1990. „An Approach to Discourse Markers." *Journal of Pragmatics* 14 (3): 383-398.

Furkó, Bálint Péter. 2011. „Diskurzusjelölők és szövegösszefüggés: a kohézió vagy a koherencia eszközei? Az implikáció mint a szövegösszefüggés eszköze." In *A szövegösszefüggés elméleti és gyakorlati megközelítési módjai: poliglott terminológiai és fogalmi áttekintés : Petőfi S. János 80. születésnapjára* [Theoretical and empirical approaches to coherence. A polyglott overview of terminology and concepts. In honour of the eighty-year-old János S. Petőfi], edited by Edit Dobi, 37-55. Debrecen: Debreceni Egyetemi Kiadó.

Ho-Dac, Lydia-Mai and Catherine Bolly. 2011. „Approches fonctionnelles de la structuration des textes." *Discours* (8).

Ho-Dac, Lydia-Mai, Julie Lemarié, Marie-Paule Péry-Woodley, and Marianne Vergez-Couret. 2012. „Multidisciplinary Perspectives on Signalling Text Organisation." *Discours* (10).

Luscher, Jean-Marc. 1994. „Marques de connexion : procédures de traitement et guidage inférentiel." In *Langage et pertinence. Anaphores, connecteurs, temporalité, lexique*, edited by Jacques Moeschler, Anne Reboul, Jean-Marc Luscher and Jacques Jayez, 175-227. Nancy: Presses Universitaires de Nancy.

Mel'cuk, Igor A., André Clas, and Alain Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot.

Petőfi, János S.. 1996. *Az explicitség biztosításának feltételei és lehetőségei természetes nyelvi szövegek interpretációjában* [Conditions and possibilities of ensuring expliciteness in the interpretation of natural language texts]. Budapest: MTA Nyelvtudományi Intézet.

Petőfi, S.János. 2004. *A szöveg mint komplex jel*. Budapest: Akadémiai Kiadó.

Petőfi, János S. and Zsuzsa Benkes. 1998. *A szöveg megközelítései*. Iskolakultúra Vol. 2. [Approaches to text]. Budapest: Iskolakultúra.

Redeker, Gisela. 2006. „Discourse Markers as Attentional Cues at Discourse Transitions." In *Approaches to Discourse Particles*, edited by Fischer Kerstin, 339-358. Amsterdam: Elsevier.

Rudolph, E. 1996. *Contrast: Adversative and Concessive Relations and their Expressions in English, German, Spanish, Portuguese on Sentence and Text Level*. Berlin - New York: Walter de Gruyter.

Schiffrin, D. 1988. *Discourse Markers*. Cambridge: Cambridge University Press.

Corpus resources

Compleat Lexical Tutor v.6.2, Cobb, Tom, Université du Québec à Montréal, http://www.lextutor.ca/

Agenda 2000 - Volume I - Communication : FOR A STRONGER AND WIDER UNION. DOC/97/6, Strasbourg, 15 July 1997, http://europa.eu/rapid/press-release_DOC-97-6_en.htm?locale=en

Agenda 2000 - Volume I - Communication :POUR UNE UNION PLUS FORTE ET PLUS LARGE. DOC/97/6, Strasbourg/Bruxelles, le 15 juillet 1997, http://europa.eu/rapid/press-release_DOC-97-6_fr.htm?locale=en

Maupassant, Guy de: Bel Ami. Association de Bibliophiles Universels, http://abu.cnam.fr/

Maupassant, Guy de: Bel Ami. eBooks@Adelaide, The University of Adelaide Library, http://ebooks.adelaide.edu.au/m/maupassant/guy/m45bl/index.html

6.
Semiotic-textological approaches to literary discourse[1]
KÁROLY ISTVÁN BODA — JUDIT PORKOLÁB

In this study, we would like to select and review the main results of a few studies published in the previous volumes of the periodical *Officina Textologica* which, according to their specific research aims and objectives, carried out semiotic textological analysis of literary texts using or examining the approach of co-referential analysis. Our main concern is to overview the basic methods and formalism of co-reference analysis which have been gradually developed from the first volume of *Officina Textologica* and become a powerful tool to explore the textological structure and thematic composition of literary texts.

1. The theoretical background: János S. Petőfi's conception about the polyglot research program in textology / text linguistics realized in *Officina Textologica*

1.1

In the first volume of Officina Textologica, János S. Petőfi outlined a polyglot reseach program in textology and text linguistics which, since the first volume in 1997, has been thoroughly elaborated throughout the studies and essays of various researchers in the subsequent volumes of the periodical.

According to the author, "the linguistic-based research on texts has been gradually intensified since the 1960's. Because of the linguistic background, this particular field of research has been referred to as *text linguistics* on the one hand, and *text theory*, on the other hand (the latter emphasizing the theoretical nature of the research). In addition, we can find terms which contain 'discourse' instead of 'text' forming 'twin pairs' of the same term", i.e. to refer more or less to the same field of research (Petőfi 1997: 7). As regards the terminology used in the periodical *Officina Textologica*, both the term 'textology' and 'text linguistics' suggest the use of traditional language-based linguistic tools, but **the term 'textology' also involves the use of specific 'textological' tools, i.e. tools which can operate with the knowledge about the world**. It is especially important when we analyse the semantic aspects of texts. In this sense, the term 'text linguistics' can be considered as the linguistic component of textology.

In the first chapter entitled 'The disciplinary environment of text study. Text linguistics and textology in text research', the author outlines the concept of 'textology' and 'text linguistics' in terms of their relationship with the other fields of research (see Fig 1).

In Fig 1, we can see the fields of research related to textology, the applications of textology, as well as the relationships of the various fields and applications illustrated here. "Because the textology of verbal texts in language L requires the text linguistics of verbal language L, which itself is built on the linguistics of the systems of verbal language L, and the linguistics of the use of the elements of those systems, the text linguistics of verbal language L provides a connecting link between the non-text-linguistic sector of linguistics (which, all things considered, is a sector of the grammar of system sentences), and the sector of textology" (ibid. 11). In Fig 1, the terms 'verbal texts' and 'verbal language L' refer to the fact that similar disciplinary environments can be (and should be) created for multimedia texts in the same way.

APPLICATIONS OF THE RESEARCH RESULTS OF TEXTOLOGY AND RELATED FIELDS, AND LINGUISTICS		
the domain of various types of texts: everyday texts, scientific, legal, religious, political, literary, etc. texts		
RELATED FIELDS TO TEXTOLOGY OF VERBAL TEXTS IN LANGUAGE "L" **poetry, narrative, rhetoric, stylistics**	TEXTOLOGY OF VERBAL TEXTS IN LANGUAGE "L"	LINGUISTICS OF VERBAL LANGUAGE "L" • **linguistics of the (lexical and prosodic-rhythmic) system of verbal language "L", and linguistics of the use of the elements of its system** • **text linguistics of verbal language "L"**
GENERAL POETRY, NARRATIVE, RHETORIC, STYLISTICS	GENERAL TEXTOLOGY	GENERAL SYSTEM AND TEXT LINGUISTICS OF VERBAL LANGUAGES
DISCIPLINES OF THE INTERDISCIPLINARY FOUNDATIONS OF TEXTOLOGY **philosophy, psychology, sociology / anthropology, semiotics, communication theory, disciplines of formal and empirical methodologies**		

Figure 1: The Disciplinary Environment of Text Study (abbrev.) (Petőfi 1997: 10)

It can be clearly seen from Fig 1 that a possible application of textology, and therefore *semiotic textology* as well, is the analysis of literary texts.

In the first volume of *Officina Textologica*, János S. Petőfi deals primarily with two major features of textuality: co-referential elements and co-reference relations, and the linear arrangement of text constituents.

1.2

In this study our main concern is the review of the possible applications of co-referential elements and co-reference relations in the analysis of literary discourse (the issue of the possible linear arrangements of text constituents is thoroughly discussed in another study of this volume). Therefore, it is worth outlining the main concepts and ideas about co-referentiality expounded in the first volume of *Officina Textologica* by János S. Petőfi. First let us see a brief definition of the terms: "It is well-known that **co-referential elements** are referred to as those language elements or phrases which might be different in their verbal manifestations but, according to the conviction of the reader or interpreter, they **all refer to the same entity or fact of the world the text describes**. According to this sense, co-reference relations are referred to as the relations between the supposed co-referential elements." (Petőfi 1997: 24)

In the first volume, several examples illustrate the co-referential elements and co-reference relations within selected texts. The first example is a passage from an essay by *Gyula Illés* about the famous Hungarian poet, *Endre Ady*. The essay is entitled '*About Ady's heritage*' (ibid. 25). In this passage, the chain of elements which have co-reference relation with the proper noun 'Ady' can be relatively easily revealed taking almost exclusively the lexical-syntactic relationships into account. In the example below, we illustrate the corresponding co-referential elements (in our case nouns, pronouns etc.) by underlining them. The text sentences are denoted by ordinal numbers in square brackets:

> [K00] About <u>Ady</u>'s heritage (extract)[2]
> [K01] <u>He</u> had an impact. [K02] <u>His</u> individual and exceptional ability to compose sentences, imagery, expressions, etc. has been imitated merely by mediocrity, but <u>he</u> was indeed the one <u>who</u> teached the next generation those things <u>he</u> had also inherited from <u>his</u> predecessors: the poet's courage and boldness to take on new thoughts and ideas, and to face the challenges. [...]

The second example is a passage from a short prose by *Lajos Áprily* entitled '*Birds above the sea*' (ibid. 26).

---

[2] (translated by BIK)

[K00] Birds above the sea (extract)[3]

[K01] On the bridge of the Danube, there are people standing near the bars, adults as well as children. [K02] They are looking at the gulls swinging above the river and, from time to time, landing on the water. [K03] In our reach of the Danube, too, I often see those birds with coral red beaks hovering over the water. [K04] Sometimes they fly high, and my eyes are loosing sight of them in the blueness or in the mist.

[K05] But I have seen gulls above the sea, too. [...]

In order to illustrate the co-reference relations in a more general sense, in this example János S. Petőfi takes into consideration the following four groups of co-referential elements:

0. conjunctions (including adverbial conjunctions of time)
1. elements which refer to the narrative self ("I")
2. elements which refer to the birds
3. elements which refer to the Danube and the sea.

Table 1 shows in its columns some co-referential elements of the passage belonging to the corresponding groups given above. As regards the question which elements belong to a particular group, it is important János S. Petőfi's remark (given in the fourth chapter) that the relationship of the elements within a group (e.g. birds, gulls, beaks etc.) "can only be explicitly established using a lexicon (i.e. a vocabulary having a specific structure) which also contains a thesauristic arrangement of its entries" (Petőfi 1997: 54)

text sentences	types of co-referential elements			
	0.	1.	2.	3.
[K00] Birds above the sea			**birds**	the **sea**
(end of paragraph / title)				
[K01] On the bridge of the Danube, there are people standing near the bars, adults as well as children.				(on the bridge of) the **Danube**

---

[3] (translated by BIK)

[K02] They are looking at the gulls swinging above the river and, from time to time, landing on the water.			at the **gulls**	(above) the **river** (on) the **water**
[K03] In our reach of the Danube, too, I often see those birds with coral red beaks hovering over the water.	too often	(in) **our** (reach of the Danube), **I** (often see)	**those birds** with coral red **beaks**	in our **reach** of the **Danube** (over) the **water**
(end of paragraph)				
[K04] Sometimes they fly high, and my eyes are loosing sight of them in the blueness or in the mist.	someti mes	**my** (eyes)	**they** [⇨ gulls]	
[K05] But I have seen gulls above the sea, too.	but too	**I** (have seen)	gulls	(above) the **sea**
(...)				

Table 1: Co-reference relations in a more general sense (slightly modified) (based on Petőfi 1997: 27-31)

As regards the importance of examining the various co-referential elements and their functions, "one of the tasks of text linguistics / textology is to analyse the functions that can be performed by words from different grammatical categories and by various syntactic structures as co-referential elements to ensure textuality." However, the author also emphasizes that "textuality can always be produced by several (linguistic and non-linguistic) factors that function together" (Petőfi 1997: 31).

The author lists certain lexical categories[4] that may serve a direct or indirect function to mark (potential) co-referentiality among text sentences. In Fig 2, the examples are taken from the extract from '*Birds above the sea*' given above.

---

[4] The categories of the parts of speech listed here are based from '*Brief Hungarian grammar*' by Endre Rácz and Etel Takács. (ibid. 32)

I. 1. The verb: stand, look at, see etc.

II. Nominal words:
1. The noun: the Danube, people, gull, sea etc.
2. The adjective: coral red etc.
3. The numeral: —
4. The pronoun: they, I etc.

III. Other (independent) parts of speech:
1. The adverb: there etc.
2. The infinitive: —; the participle: swinging, landing, hovering etc.
3. The interjection: —

IV. Not independent (auxiliary) parts of speech:
1. Verbal prefix: —
2. Postposition: —; preposition: above, over etc.
3. The article: the.
4. The conjunction: and, too, but etc.
5. The modifiers: —

Figure 2 Lexical categories functioning as (potential) co-referential elements

1.3

One of the most important aspects of the analysis of literary texts is to reveal their information content, that is, the information they convey. The basic problem is that "the great majority of texts contain only as much information expressed explicitly by the language, as is just enough for us to understand (i.e. to attach meaning to) the given text on the basis of our knowledge on the language and our knowledge about the world. If texts provided in every case all relationships, concerning the given language or part of the world that is referred to, by expressing them in a lexical way, they would become »unreadable«. But it is not necessary because we, the readers of the texts, are capable, on the basis of the syntactic/semantic information provided by the text itself, to think about the given text together with the relationships necessary to understand it in an almost automatic process — at least on the literal level of interpretation." (ibid. 34)

However, when we want to examine and describe co-referential elements and co-reference relations in a given text, it can certainly be useful — and sometimes necessary — to explicitly represent the information which can be obtained or deduced from the text on the basis of our linguistic or other knowledge. In his essay, János S. Petőfi demonstrates the nature of such information by examining a narrative passage from the Bible (ibid. 32).

The Birth of Moses (extract)

[1]Now a man of the house of Levi married a Levite woman, [2]and she became pregnant and gave birth to a son. When she saw that he was a fine child, she hid him for three months. [3]But when she could hide him no longer, she got a papyrus basket for him and coated it with tar and pitch. Then she placed the child in it and put it among the reeds along the bank of the Nile. [4]His sister stood at a distance to see what would happen to him. (Ex 2:1-4)

(The Bible. New International Version, 1984.[5])

When we try to reveal the information content of the text, the basic idea is to complete the text with words having individual (lexical) meaning in order to show the lexical references which are expressed implicitly by certain words (e.g. pronouns) or grammatical structures (e.g. the subject of sentences[6]) of the text. In addition, the completed text will be more compact and much easier to analyse if we introduce the so-called *co-reference indices* which replace, with corresponding codes, the persons, things, places etc. mentioned in the text. In the short passage given above, the list of co-reference indices can be as follows:

i01 = a man of the house of Levi,

i02 = a Levite woman married by i01 = the mother of i03,

i03 = the son of i01 and i02 = the ("fine") child = Moses

i04 = the sister of i03

...

i08 = the papyrus basket

i09 = the reeds

i10 = the Nile

i11 = the bank of i10

i12 = (at) a distance (from i11)

...

In addition to the introduction of co-reference indices, we shall use the concatenation (^) sign to mark the sequence ("concatenation") of words belonging to the same expression having an individual co-reference index. Moreover, we shall place in brackets those words and their indices which do not actually appear in the text but we insert them in order to reveal all the information that the text contains.

The Birth of Moses[i03] (completed with co-reference indices)

[1]Now   a^man^of^the^house^of^Levi[i01]   married

---

a^Levite^woman[i02], [2]and she[i02] became pregnant and (she[i02]) gave birth to a^son[i03]. When she[i02] saw that he[i03] was a^fine^child[i03], she[i02] hid him[i03] for three months. [3]But when she[i02] could hide him[i03] no longer, she[i02] got a^papyrus^basket[i08] for him[i03] and coated it[i08] with tar and pitch. Then she[i02] placed the^child[i03] in it[i08] and put it[i08] among the^reeds[i09] along the^bank[i11]^of^the^Nile[i10]. [4]His[i02]^sister[i04] stood at a^distance[i12]^(from^the^bank[i11]) to (she[i04]) see what would happen to him[i03]. (Ex 2:1-4)

The information we have added to the actual lexical manifestation of the given text as a form of co-reference indices ("replacements") can be deduced from purely grammatical knowledge. "These replacements can be performed because

- when we start reading a text, in our mind (technically speaking, in our mental model) the story becomes vivid: its entities (characters, objects, places etc.) appear and 'begin to live';
- in their first appearance, the entities are usually referred to by words or phrases having individual (lexical) meaning. Later, when they appear again, they might be referred to by other words, e.g. pronouns or determiners (or affixes in Hungarian, etc.); we are capable to establish the co-reference relationship between them and the words with individual meaning the entities have been first referred to by.

During the reading (or hearing) process we are always thinking in two dimension: first, in the dimension of text (looking for the relations between language elements without individual meaning and words with individual meaning); second, in the dimension of the relationship between the text and our mental model (looking for the relations between the referential elements of the text and the people, things, places, etc. in the mental model)." (ibid 33-34.)

Because our main concern here is about the analysis of literary discourse, we would like to make two important remarks which show the interpretation power of co-reference analysis.

First, in some cases grammatical (and/or linguistic) knowledge is not enough and we need knowledge about the world to complete the text with the necessary information. For example, on the basis of the very existence of Moses' sister (i04) who, in the Biblical passage, watched Moses when he was in the basket, we have to suppose that Moses was not the first child of his parents. Although it is obvious from our basic knowledge about the world, there is no hint in the passage that she (or any sister / brother of his) was born before Moses. But our knowledge is adequate — actually, it is well known from (the other parts of) the Bible that Moses had an older sister, Miriam, and an older brother, Aaron.

Second, a very important aspect of the selected passage from the Bible is that there are some elements *which cannot be understood from the text alone*. Our common knowledge (which is based on our everyday experiences) says that a sane woman will never get rid of her newborn child — unless she is compelled by extraordinary circumstances. This contradiction should be resolved by revealing what could have happened; it is the very point of the analysis when simple (i.e. intratextual) text linguistic methods are not sufficient any more and *intertextuality* is required. In fact, from the previous book of *Exodus* we can learn that the Egyptians were so afraid of the Israelites that "Pharaoh gave this order to all his people: »Every boy that is born [to the Hebrews] you must throw into the Nile, but let every girl live. «" (Ex 1:22)[7]

In the three examples presented above János S. Petőfi introduced three analytical approaches to illustrate co-referentiality in texts:

- the first example illustrated a single chain of co-referential elements underlined within the text;
- in the second example four different chain (or group) of co-referential elements were listed in separate columns of a table which itself represented the relationship between (the group of) co-referential elements and text sentences;
- in the third example all persons, objects and events which occurred (i.e. which were referred to) in the text were formally distinguished and represented by co-reference indices.

But, as the author says, "none of these examples can be considered an exact text linguistic / textological analysis because they lack the most important factor without which the analysis cannot be complete: to determine, on the basis of a thorough examination, whether using (text) linguistic methods alone is enough to explore co-referentiality or we should employ other — that is, textological! — methods as well" (ibid. 37-38).

2. The gradual development of the methods and formalism of co-reference analysis and its applications to the analysis of literary texts

2.1

In the second volume of *Officina Textologica*, giving a normative sample to the other authors of the volume, János S. Petőfi introduced and explained the basic formalism of co-reference analysis by selecting and analysing another passage from the Bible.

---

[7] The Bible. New International Version, 1984.
(http://niv.scripturetext.com/exodus/1.htm, 2012-04-25)

> **Ve:[8]** The Calling of Matthew (Mt 9:9-13)
>
> [9]As Jesus went on from there, he saw a man named Matthew sitting at the tax collector's booth. "Follow me," he told him, and Matthew got up and followed him. [10]While Jesus was having dinner at Matthew's house, many tax collectors and "sinners" came and ate with him and his disciples. [11]When the Pharisees saw this, they asked his disciples, "Why does your teacher eat with tax collectors and 'sinners'?" [12]On hearing this, Jesus said, "It is not the healthy who need a doctor, but the sick. [13]But go and learn what this means: 'I desire mercy, not sacrifice.'[9] For I have not come to call the righteous, but sinners."
>
> (The Bible. New International Version, 1984.[10])

The main steps of co-reference analysis performed by János S. Petőfi can be summarized as follows (Petőfi 1998: 15-17):[11]

- the denotation of text sentences by **[Kxx]** (printed in bold type), where xx is the ordinal number of the corresponding text sentence;[12]
- the insertion of *additional information*, based on the text sentences themselves, their verbal context and/or our knowledge about the world, into the text sentences in order that they should be considered as individual information units. The inserted verbal phrases are presented in brackets and in italic (cursive) type;
    - when a verbal phrase is considered as one unit, its constituents are linked together by the concatenation (^) sign;
- the detailed explanation of the syntactic and co-referential analysis of text sentences is referred to as 'Commentaries' (not presented here but can be found in the complete co-reference analysis of the original study);

---

[8] 'Ve' is the abbreviation for 'Vehiculum' referring here to the selected passage to be analysed.

[9] „For I desire mercy, not sacrifice, and acknowledgment of God rather than burnt offerings." (Hosea 6:6)

[10] http://niv.scripturetext.com/matthew/9-9.htm, 2012-04-25.

[11] A more detailed description of the main steps of co-reference analysis can be found in another study of this volume by *Andrea Nagy* and *Franciska Skutta*.

[12] A text sentence is (by definition) begins with a capital letter and ends with a full stop (.), colon (:), exclamation mark (!) or question mark (?) (Petőfi 1998: 15). However, we would like to follow the steps of the original analysis of the selected passage presented by János S. Petőfi where the language of the passage was Hungarian. Because there are differences in punctuation in the Hungarian and English translations of the selected passage, in some cases we will not stick to the above definition when distinguishing text sentences.

- the list and description of *co-reference indices*; they are denoted by [ixx] where xx is the ordinal number of the corresponding index;
- the insertion of the co-reference indices into the text sentences,
  - placing an index immediately after the verbal phrase referred to by the index,
  - in case an index denotes a nominal phrase, its first occurrence is denoted by an equals sign (=) and in bold typeface.

The author remarks that "for the sake of simplicity, I will not deal with the analysis of articles, negative adjectives/adverbs, adverbial phrases of time, verb tenses and conjunctions" (Petőfi 1998: 16).

Now let us see first the co-reference analysis of the selected passage from the Bible, and then the list of the corresponding co-reference indices (cf. Petőfi 1998).

Ve/&ind:[13]
[K00] The Calling(i03)[i00] of Matthew[i03] (by Jesus[i01])
[K01] As Jesus[=i01] went^on[i01] from there[=i02], he[i01] saw[i01|i03] a^man^named^Matthew[=i03] (who[i03] was^)sitting[i03] at the^tax^collector's^booth[=i04].[14]
[K02] "(You[i03]) Follow[i03|i01] me[i01],"
[K03] he[i01] told[i01|"K02"] him[i03],
[K04] and Matthew[i03] got^up[i03] and (Matthew[i03]) followed[i03|i01] him[i01].
[K05] While Jesus[i01] was^having[i01|i09] dinner[=i09] at Matthew[i03]'s^house(i03)[=i05], many^tax^collectors[=i06] and (many^)"sinners"[=i07] came[i06,i07] (to the^house[i05]) and (they[i06,i07]) ate[i06,i07] with him[i01] and his[i01]^disciples(i01)[=i08] (in[15] the^house[i05]).
[K06] When the^Pharisees[=i10] saw[i10|"K05"[16]] this[K05], they[i10] asked[i10|"K07"] his[i01]^disciples(i01)[i08],
[K07] "Why does your[i08]^teacher(i08)[i01] eat[i01] with

---

[13] 'ind' is the abbreviation for 'indexed' referring here to the fact that the vehiculum has been analysed, indexed and completed by co-reference indices.
[14] Introducing a new co-reference index for 'tax collector' in general, a partial co-reference could be established with i06 (and with i03, etc.). Note that in the Hungarian translation the word 'custom' occurs instead of the 'tax collector's booth'.
[15] or (more probably) *in front of* the house (see later)
[16] János S. Petőfi's remark emphasizes that [K05] here "refers to 'a more general event' than 'this' in the given text sentence", but, for the sake of simplicity, we do not introduce another index to denote the corresponding part of [K05] which has been actually referred to (cf. Petőfi 1998: 23).

tax^collectors[i06] and 'sinners'[i07]?"[17]

[K08] On (Jesus[i01]) hearing[i01|"K07"] this[K07], Jesus[i01] said[i01|"K09^K10^K11^K12"] (to the^Pharisees[i10]),

[K09] "It is not the^healthy(^men)[=i11] who[i11] need[i11|i13] a^doctor[=i13], but the^sick(^men)[=i12] (need[i12|i13] a^doctor[i13]).

[K10] But (you[i10]) go[i10] and (you[i10]) learn[i10|i14] what[=i17] this[K11^K12] means["K11^K12"|i17]:

[K11] 'I[i01] desire[i01|i14] mercy[=i14], (I[i01] do^)not (desire[i01|i15]) sacrifice[=i15].'

[K12] For I[i01] have^not^come[i01] to (I[i01]) call[i01|i16] the^righteous(^men)[=i16], but (I[i01] have^come[i01] to (I[i01]) call[i01|i07] the^)sinners[i07]."

The list of co-reference indices that have been used is as follows:

i00the calling of Matthew[i03] by Jesus[i01] (did not occur in the original
analysis)

i01Jesus

i02the place where Jesus[i01] was before the events of the passage

i03Matthew

i04the tax collector's booth

i05Matthew[i03]'s house

i06many tax collectors (that came to (i03)[i05])

i07many "sinners" (that came to (i03)[i05])

i08Jesus[i01]'s disciples

i09dinner (at (i03)[i05])

i10the Pharisees

i11the healthy (men), *in general*

i12the sick (men), *in general*

i13a doctor, *in general*

i14mercy, *in general*

i15sacrifice, *in general*

i16the righteous (men), *in general*

i17the meaning of [K11] and [K12] (not occurred in the original analysis)

As we can see, there are some additional rules which have been applied in the formal description of the text under analysis:

---

[17] Although in [K07] the phrase "taxt collectors and sinners" (and later, in [K12] the word "sinners") refers to tax collectors and sinners *in general* (i.e. to any of such persons), the author preferred the use of existing indices [i06] and [i07] (and [i07] in [K12], respectively). (cf. Petőfi 1998: 24)

(1) the co-reference index which denotes the *subject* of a particular verb should be placed right after the verb;

(2) the co-reference index which denotes the *direct object* of the verb is placed after the index denoting the subject, and the indices are separated by the vertical line (|) sign;

(3) when the subject (or the direct object) consists of more than one indexes (i.e. it can be expressed by a set of indexes), those co-reference indexes are separated by the comma (,) sign; when the subject or the direct object consist of more than one index of text sentences, those indices are separated by the concatenation (^) sign;

(4) the *possessive noun* should be placed before its possession in brackets.

It is very illuminating and thought-provoking to overview the occurrences of co-referential elements in text sentences illustrated in Table 2 (cf. ibid. 29-30). (For the sake of simplicity the verbs are displayed at those rows only which belong to the subject of the corresponding verb).

Although the author does not deal with the analysis of the text as literary discourse, it can quite possibly be stated that the co-reference analysis of the text reveals details that no other analytic method is likely to reveal. For example, it is worth observing that in the selected passage there are at least three levels of meaning:

(1) the description of the events which actually happened at the time of Matthew's calling;

(2) the communication between first Jesus and Matthew, then between the Pharisees and Jesus. Note that the indices of the quoted sentences, which are embedded into the analysed passage, occur in the same grammatical position where we otherwise use co-reference indices referring to "real" entities of the first level;

(3) the general (or abstract, symbolic, etc.) meaning of all (or some) of the events and related behaviour which have been transformed, by the questions of the Pharisees and the answers of Jesus, into the ethical rules of behaviour to be followed (including the 'right' or 'wrong' assessment of them).

Moreover, there is one point in the selected passage which, in case we perform only an intuitive analysis, is very easy to miss. It is the (quite unexpected) appearance of the Pharisees which does require an explanation. The necessary knowledge, when our basic knowledge about the world is not enough, might come from some *background knowledge,* e.g. from other texts. In our case, a significant text sentence from Luke can help: "Then Levi held a great banquet for Jesus at his house, and a large crowd of tax collectors and others

were eating with them." (Lk 5:29)[18] In the feast a great number of people participated, so it should have been outside the house of Levi (Matthew). Accordingly, the Pharisees, who were eager to examine every step of Jesus, could see the event, so they could make their comments.

---

[18] The Bible. New International Version, 1984. (http://niv.scripturetext.com/luke/5-29.htm, 2012-05-05)

	K01	K02	K03	K04	K05	K06	K07	K08	K09	K10	K11	K12
i01	Jesus went^on ──── he saw	me	he told	him	Jesus was^ having ──── him his^ disciples	his^ disciples	your^ teacher eat[i01]	(Jesus) hearing ──── Jesus said			I desire ──── I not (desire)	I have^ not-^come ──── (I) call ──── (I) (have^ come) ──── (I) (call)
i02	there											
i03	a^man^ named^ Matthew ──── (who) (was^) sitting	(You) Follow	him	Matthew got^up ──── (Matthew) followed	Matthew's ^ house							
i04	the^tax^ collector's-^booth											

	K01	K02	K03	K04	K05	K06	K07	K08	K09	K10	K11	K12
i05					Matthew's^ house the^house the^house							
i06					many^tax-^collectors came (they) ate		tax^ collectors					
i07					(many^)"sin ners" came (they) ate		'sinners'					(the^) sinners
i08					his^ disciples	his^ disciples	your					
i09					dinner							
i10						the^ Pharisees saw they asked		the^ Pharisees		(you) go (you) learn		

	K01	K02	K03	K04	K05	K06	K07	K08	K09	K10	K11	K12
i11									the^ healthy(^men) need			
i12									the^sick-(^men) (need)			
i13									a^doctor ——— (a^doctor)			
i14											mercy	
i15											sacrifice	
i16												the^ righteous-(^men)
i17										what		

2.2

In addition to János S. Petőfi's study, there are several studies in the second volume of *Officina Textologica* which selected a literary text to be analysed following János S. Petőfi's sample analysis. We would like to overview their contributions to the development of the formalism of co-reference analysis as well as the comments of *János S. Petőfi* and *Edit Dobi* (Petőfi-Dobi 1998) who provided a detailed discussion of all the studies published in the volume.

In their study, *Károly I. Boda* and *Judit Porkoláb* (Boda-Porkoláb 1998) selected a passage from the Revelation (Rev 21:9–23). Although the text is very rich in *intertextual references*, some of them are deliberately ignored because they do not have explicit co-reference relation to the other sections of the passage. However, the authors select one of such references ("the seven bowls full of the seven last plagues") and attach a co-referential index to it to demonstrate its importance. "The entity that belongs to the phrase 'the seven bowls full of the seven last plagues' does not occur in the selected passage any more; nevertheless, its selection as a co-reference index can be accounted for by the fact that it makes a reference to the antecedents of the selected passage (i.e. to the sequence of visions describing the final judgement in the Revelation)." (ibid. 37) The authors raise the problem that some of the co-reference relations could be lost if we indicate only the co-reference index of the subject and that of the direct object after the predicate (e.g. in the verb pattern sb_1 *tell* sb_2 *sth* the indirect object of the verb 'sb_2', although not indicated, might be important when referring to a person in the text) (ibid. 37). Another interesting aspect of the selected passage from the Bible is that it can be interpreted using different paradigms each having a different level of meaning which might (perhaps should) be reflected in the interpretation of the corresponding co-reference indices; e.g. New Jerusalem could be interpreted as the ancient Jewish city, the Church of Jesus, the human soul or God's Kingdom (ibid. 40).

In her study, *Márta Tuba* (Tuba 1998) selected a folk (*Palots*) tale for analysis. The author introduced several new denotations; they are as follows (Tuba 1998: 142-143):

- the possessive case is represented by connecting the possession and the possessor with the plus (+) sign (e.g. in the title of the tale, the^cock**[=i01]**'s^gold^coin**[=i02**+i01]);
- a change in the meaning of a co-reference index is indicated by the asterisk (*) sign (e.g. [i04] means the king as a person in the phrase *the king's brother*, compared with [i04*] which means his role or rank in the sentence *her brother was a king*);
- a co-reference is marked with an apostrophe (') when it is used to address somebody (e.g. i01 means 'the cock', i01' means the addressing

of the cock in the sentence 'My little cock, give me the gold coin!'). Note that "the referential value of addressing is a kind of 'language reference' which is different from the real world entity which might be referred to as the addressee" (Petőfi-Dobi 1998: 252).

It is worth noticing that "the most general organizing factor of the tale is the repetition" of various segments of the text (Tuba 1998: 161). For the exact identification of those segments the author uses the denotation K18a, K18b, etc. to mark the different clauses of the text sentence K18. To emphasize the importance of the indexing of time, the author remarks that "the most characteristic feature of the narrative type of texts is the sequence of events in time, which has been indicated by a unified index i00 to mark the different adverbial phrases of time" in the analysis; for example, 'once^upon^a^time[i00]', then[i00], in^three^days[i00], etc. (ibid. 161). This index does really express the dimension of time but "does not represent the progress of time" (Petőfi-Dobi 1998:259).

In his study, *Imre Békési* emphasizes that "if the sufficient number of analyses have been performed and their descriptive apparatus is unified to the extent that their results can be coded for computer processing, then we can seriously hope for new discoveries" (Békési 1998: 185). His starting point is the so-called 'thesis sentence' which, in some type of texts, comprises (either before or after, and in rare cases in the middle of, the corresponding sentences) the essence of the content that the text describes. In the selected passage from *Anna Karenina* by Leo Tolstoy, the first sentence can be considered as a thesis sentence which contains all the co-reference indices which occur in the following text sentences of the passage: "The mother cannot even compare Vronsky with Levin (*because she thinks the characters of the two men so different*)." Note that the completion of the text sentence with explanatory information (given in brackets) seems to be very important because the organization of text sentences "is realized under the control of contextual-logical (semantic) relations as well as communication, rhetoric etc. rules" (ibid. 191). As a consequence, the co-reference relations of the text sentences of the selected passage can be organized (or illustrated) by the use of conjunctions such as 'because', 'therefore', and 'however' (ibid. 192-193).

In her study, *Katalin B. Fejes* analyses the poem '*Greeting Thomas Mann*' by the famous Hungarian poet, *Attila József*. Her approach to the poem is determined by the fact that the poem "is actually a greeting, therefore its creation could have been affected by the conceived situation where the poem was to be read out loud. This situation could have been so effective that its influence was imprinted on the poem." (B. Fejes 1998: 194) According to the main senses, the author distinguishes three different communication channels each containing a corresponding chain (or "bundle") of co-referenced lexemes from the poem:

auditive channel	kinetic channel	visual channel
[K01] tell a tale (whether he wishes) the tale tell a tale [K02] Tell us (what ...) tell a tale (about ...) [K03] tell us the truth [K05] your gentle voice tell a tale (about ...) [K06] you can read out loud [K07] if you speak [K09] let's start telling a tale [K10] We are listening to you	[K01] Please, don't go (or he wishes) you to be there Sit down among us [K02] you are here with us [K08] Please take a seat	[K10] who is just looking at you (who) sees a European

Table 2: Three communication channels and the co-referenced lexemes belonging to them (ibid. 195; slightly modified)

The author assigns co-reference indices to (some of) the nominal and (some of) the verbal elements of the selected text, as well as "to the referential values of the statements" (Petőfi-Dobi 1998: 260), focussing on the auditive-kinetic-visual relationships represented by the given channels. These relationships play a central role in the interpretation of the poem because "the joy of greeting can hardly be expressed by mere words, so the 'we' are expecting words from the 'you': ask, request, question, or look, rejoice, see. Therefore we might suppose

that the composition of the whole text is centered around the auditive-kinetic-visual representation of the very moment of 'we ask you'." (ibid. 205)

Interestingly enough, the composition of the poem also contains some outside references. There are some elements in the text which, although do not have explicit co-reference relationship with the other elements, refer to certain knowledge outside the textual world of the analysed text (e.g. the two main characters of *The Magic Mountain* by *Thomas Mann*, the funeral of the famous Hungarian poet, *Dezső Kosztolányi*, etc.) (ibid. 205). Although the author does not explicitly mention, to represent those references we might as well add another channel concerning memories and imagination; i.e. some "mental content" that can be associated with the situation the poem describes. This shows that the idea of identifying co-reference bundles (i.e. co-reference chains the elements of which refer to a central theme or topic of the text) can be a powerful tool for the interpretation of a literary text, fully supported by the formalism of co-reference analysis.

As regards the formalism used by the author, in the co-reference structure of the analysed text we can observe "full" as well as partial or thesauristic co-reference relationships between the corresponding lexemes within each bundle (e.g. the first occurrence of '*tell a tale*' in K01, where a child asks its parent(s), has a partial reference to '*tell a tale*' in K01 / K02 etc. where the 'we' ask the 'you' in the textual world of the poem; note that the subsequent occurrences of '*tell a tale*' are in "full" co-reference relation with each other referring to the same general entity and/or action; an example for the thesauristic co-reference relationship might be 'you *speak*' vs. 'we are *listening* to you'). Note that the issue of the identification and/or denotation of the partial co-reference relationship also arises in other studies (e.g. in Tolcsvai 1998: 229, Petőfi-Dobi 1998: 261, etc.).

In his study, *Gábor Tolcsvai Nagy* analyses a selected part of a so-called 'dictionary' novel "*Dust*" by *Ferenc Temesi*. The selected text has a form of a dictionary entry having the headword '*ideal*' (or '*Dust.ideal*' to indicate the place of the 'headword' within the 'dictionary'). It includes several references which establish connection with the other parts of the dictionary novel and therefore can only be fully understood with the comprehensive knowledge of the full text. The co-reference analysis made by the author, following the basic methodology and formalism of János S. Petőfi's normative sample (Petőfi 1998), is intended to be an element of "developing a textological representation language" (Tolcsvai 1998: 223).

The analysis has several aspects that are worth considering. As an example, let us see the co-reference analysis of the fourth text sentence where the

narrative self of the selected text, who in his youth[19], 'literally' chose Percy Harrison Fawcett as his ideal, expresses his deep emotions:

> [K04/&ind] In it,[20] this valiant^British^traveller[i06] related his[i06] adventures in South America and he[i06] suffered so many sore trials in the jungle of Amazonas that, considering Zola or not, it [he? they?; see the 'commentary' below] really touched[i08|i03] me deeply.[21]

i03the narrative self, "I"
i06Percy Harrison Fawcett, the "valiant British traveler"
i08the destiny or behaviour of Percy Harrison Fawcett

Note that several co-reference indices are deliberately missing. As the author says, "the analysis is not 'complete' because it focuses, in the first place, on the most uncommon nominal parts, verbal arguments and pronominal elements which need to be explicated and have co-reference relation" with other elements in the selected text (Tolcsvai 1998: 225). One of the most interesting aspects of the analysis of the text sentence K04 is that the subject of the last clause ('it') is "not entirely unambiguous", partly because the original Hungarian text allows other translations as well (e.g. 'he' or 'they'). We might assume, however, that the entity that the subject refers to is, most likely, the destiny or behaviour of the traveller. "According to this interpretation, the subject of 'touched' is in partial co-reference relation with the 'adventures' and 'trials' constituents of the text sentence K04" (ibid. 229).[22] This calls our attention to the significance of the phenomenon of partial co-reference relationship which has also been mentioned in the overall discussion of the volume (cf. Petőfi-Dobi 1998: 261).

2.3

The *Epilogue* of the 2nd volume of *Officina Textologica* is an overall discussion and analysis of the main concepts, problems and further questions of co-reference analysis. Reflecting on some issues raised in the studies of the same

---

[19] as we learn from the first text sentence of the analysed text,

[20] That is, in the book first mentioned in the second text sentence of the analysed text („I got his book in Christmas"). Although it has a separate co-reference index (i07) but the author does not use it in the co-reference analysis of the fourth text sentence.

[21] (translated by BIK)

[22] Note that we can reasonably assume that one of the most important feature of an 'ideal' is his or her *behaviour* in certain, most likely difficult, situation. In this case, i08 (=the destiny or *behaviour* of Percy Harrison Fawcett) would establish a partial co-reference relation with the title or 'headword' of the 'dictionary entry' text ('ideal') which, by definition, should be the most significant word (a keyword or a kind of 'thesis word', cf. Békési 1998) of the selected text.

volume, *János S. Petőfi* and *Edit Dobi* overview and summarize the basic thoughts and questions of co-reference analysis as follows:

Co-reference analysis is performed by considering the relationship of two different levels (see Petőfi-Dobi 1998: 238):

- the first level is "the mental image of a particular segment of the world which has been expressed, according to the interpreter, in the text analysed"; as a consequence, this level contains the entities of the given world segment (i.e. persons, objects, concepts, attributes, statements etc.) and their corresponding mental images;

- the second level is "the vehiculum (i.e. the physical appearance) of the text analysed including the constituents of the text to be analysed semantically / syntactically by the interpreter. Some of these constituents refer to the entities of the represented world segment, while the other constituents express certain statements on the referred entities.

As regards the interpretation of co-reference relation, it is "a kind of relationship interpreted in some way between text constituents referring to the (supposedly) same entity of the represented world segment" (ibid.). Various types of this relationship can be distinguished, e.g. we can speak of 'identity' relationship in case all constituents refer to the same entity, or of 'possessor-possession' relationship when one referred entity (e.g. an adventure, destiny, behaviour etc.) belongs to the other referred entity (e.g. a traveller). Naturally, "there can be other types of relationship, too" (ibid.).

The authors list four very important questions to be answered when discussing 'co-reference analysis' in general:

(1) "what kind of entities can be referred to by the constituents of text";

(2) "what kind of text constituents can have reference" to the above entities;

(3) "what types of co-referentiality can be distinguished, that is, what kind of referential difference" between the referred entities can be accepted so that their relationship can still be considered co-referential;

(4) "what can be considered as the most effective way to denote co-reference relations" (ibid. 238).

As to the detailed discussion of and possible answers to these questions we refer to the chapter of this volume written by Andrea Nagy and Franciska Skutta.

As the review and discussion of the main results of the above studies presented here most probably show, the formalism and methodology of co-reference analysis can be a very effective and highly analytic starting point to the semiotic-textological analysis of texts which in turn makes a substantial and valuable contribution to the analysis of literary discourse as well.

3. Further readings: a selected bibliography

In addition to the studies reviewed before, we have compiled a short bibliography of studies analysing literary texts using (or referring to) the concepts and methods of co-referential analysis in the subsequent volumes of Officina Textologica. They are as follows:

Bencze Lóránt: A koreferáló elemek és viszonyok retorikai-stilisztikai megvilágítása. [Rhetorical-stylistic approach to coreferential elements and co-reference relations.] Officina Textologica 4/2000, 35–56.

Benkes Zsuzsa: Kreatív gyakorlatok a koreferenciarelációk analitikus elemzésének előkészítéséhez. [Creative exercises as preparation for the analytical examination of co-reference relations.] (Példaszöveg: Biblia. Mózes II. könyve. 2. rész 1-11.) Officina Textologica 4/2000, 57–80.

Boda I. Károly, Bodáné Porkoláb Judit: A koreferencia kérdései a számítógépes szövegfeldolgozás szempontjából. [Problems of co-reference in computer-based text processing.] Officina Textologica 4/2000, 150–180.

Boda I. Károly, Porkoláb Judit: Téma-réma kapcsolatok vizsgálata egy kiválasztott versszövegben korreferenciaelemzés segítségével. [Examination of theme-rheme relations in selected texts using co-reference analysis.] Officina Textologica 7/2002, 93–112.

Boda I. Károly, Bodáné Porkoláb Judit: A tematikus progresszió vizsgálata kiválasztott szövegekben korreferencia-elemzés segítségével. [Examination of thematic progression in selected texts using co-reference analysis.] Officina Textologica 9/2003, 45–58.

Boda I. Károly, Bodáné Porkoláb Judit: Egy angol vers és magyar fordításainak összevetése a korreferencialitás szempontjából. [Comparing the co-referentiality of a poem by T.S. Eliot and its two Hungarian translations.] Officina Textologica 12/2005, 83–100.

Boda I. Károly, Bodáné Porkoláb Judit: Korreferencialitás és hipertextuális kapcsolatok a poétikai kommunikációban. [Co-referentiality and hypertextual relations in poetic communication.] Officina Textologica 16/2011, 117–129.

Csűry Andrea: Korreferenciaviszonyok vizsgálata magyar és francia dialogális szövegben. [Examination of co-reference relationships in Hungarian and French dialogues.] Officina Textologica 12/2005, 19–34.

Dobi Edit: A koreferenciaelemzés kérdésének rendszernyelvészeti megközelítése szemiotikai textológiai keretben. [A system linguistic approach to co-reference analysis in semiotic textological framework.] Officina Textologica 4/2000, 81–123.

Dobi Edit: A korreferencialitás megnyilvánulásai magyar és német szövegekben. [The manifestations of co-referentiality in Hungarian and German texts.] Officina Textologica 12/2005, 101–116.

Tolcsvai-Nagy Gábor: Kérdések a koreferenciáról. [Questions about co-reference.] Officina Textologica 4/2000, 11–34.

Table 4 shows all the studies mentioned above as well as the corresponding literary text (or a certain passage of it) selected for analysis by the authors with a view to using the methods and formalism of co-reference analysis, and/or examining its main issues and problems in general.

Study from *Officina Textologica*	**Selected literary text and genre**
B. Fejes Katalin: Kompozícióalkotó korreferencia-nyalábok. [Composition-forming co-reference bundles.]	Attila József: Thomas Mann üdvözlése [Greeting Thomas Mann] (poem)
Békési Imre: A korreferenciarelációk és a tételmondat. [Co-reference relations and the thesis sentence.]	Lev Tolsztoj [Leo Tolstoy]: Anna Karenina (novel)
Bencze Lóránt: A koreferáló elemek és viszonyok retorikai-stilisztikai megvilágítása. [Rhetorical-stylistic approach to coreferential elements and co-reference relations.]	Sándor Petőfi: Akasszátok fel a királyokat! [Hang the Kings!] (poem)
Benkes Zsuzsa: Kreatív gyakorlatok a koreferenciarelációk analitikus elemzésének előkészítéséhez. [Creative exercises as preparation for the analytical examination of co-reference relations.]	The Bible. Old Testament. (Ex 2:1-10)
Boda I. Károly, B. Porkoláb Judit: Koreferenciális kifejezések és koreferenciarelációk. [Co-referential phrases and co-reference relations.]	The Bible. New Testament. (Rev 21:9-23)
Boda I. Károly, Bodáné Porkoláb Judit: A koreferencia kérdései a számítógépes szövegfeldolgozás szempontjából. [Problems of co-reference in computer-based text processing]	The Bible. New Testament. (Rev 1:1-5)
Boda I. Károly, Porkoláb Judit: Téma-réma kapcsolatok vizsgálata egy kiválasztott versszövegben korreferenciaelemzés segítségével. [Examination of theme-rheme relations in selected texts using co-reference analysis.]	Milán Füst: A szőlőműves [The Vine-dresser] (poem)

Boda I. Károly, Bodáné Porkoláb Judit: A tematikus progresszió vizsgálata kiválasztott szövegekben korreferencia-elemzés segítségével. [Examination of thematic progression in selected texts using co-reference analysis.]	Miklós Radnóti: Lomb alatt [Under the Leafy Boughs] (poem)
Boda I. Károly, Bodáné Porkoláb Judit: Egy angol vers és magyar fordításainak összevetése a korreferencialitás szempontjából. [Comparing the co-referentiality of a poem by T.S. Eliot and its two Hungarian translations.]	T.S. Eliot: Marina (poem)
Boda I. Károly, Bodáné Porkoláb Judit: Korreferencialitás és hipertextuális kapcsolatok a poétikai kommunikációban. [Co-referentiality and hypertextual relations in poetic communication.]	T.S. Eliot: The Waste Land (poem)
Csűry Andrea: Korreferenciaviszonyok vizsgálata magyar és francia dialogális szövegben. [Examination of co-reference relationships in Hungarian and French dialogues.]	János Háy: A Gézagyerek [The Géza-boy] (play)
Dobi Edit: A koreferenciaelemzés kérdésének rendszernyelvészeti megközelítése szemiotikai textológiai keretben. [A system linguistic approach to co-reference analysis in semiotic textological framework.]	Lajos Áprily: A hiúz [The Lynx] (short story)
Dobi Edit: A korreferencialitás megnyilvánulásai magyar és német szövegekben. [The manifestations of co-referentiality in Hungarian and German texts.]	Eduard Mörike: Mozart auf der Reise nach Prag [Mozart's Journey to Prague]
Petőfi S. János (1997): Egy poliglott szövegtani-szövegnyelvészeti kutatóprogram. [A polyglot textological / text linguistic research program.]	Illyés Gyula: Ady örökségéről [About Ady's heritage] (essay) Lajos Áprily: Madarak a tenger felett [Birds Above the Sea] (short prose) The Bible. Old Testament. (Ex 2:1-10)
Petőfi S. János: Koreferenciális elemek és koreferenciarelációk. [Co-referential elements and co-reference relations.]	The Bible. New Testament. (Mt 9:9-13)

	(all the texts which have been selected by the studies in the 2nd volume of Officina Textologica)
Petőfi S. János, Dobi Edit: Utószó. [Epilogue.]	(all the texts which have been selected by the studies in the 2nd volume of Officina Textologica)
Tolcsvai Nagy Gábor: Explicitség és koreferencia. [Expliciteness and co-reference.]	Ferenc Temesi: Por [Dust] (novel)
Tolcsvai-Nagy Gábor: Kérdések a koreferenciáról. [Questions about co-reference.]	Ferenc Temesi: Por [Dust] (novel)
Tuba Márta: Koreferenciális kifejezések és koreferenciarelációk egy mesében. [Co-referential phrases and co-reference relations in a folk tale.]	A kakas aranygarasa [The cock's gold coin] (folk tale)

Table 3: Selected studies from Officina Textologica and the literary texts analysed by them

References

B. Fejes Katalin (1998): Kompozícióalkotó korreferencia-nyalábok. [Composition-forming co-reference bundles.] Példaszöveg: József Attila: Thomas Mann üdvözlése. Officina Textologica 1998/2, 194–207.

Békési Imre (1998): A korreferenciarelációk és a tételmondat. [Co-reference relations and the thesis sentence.] Példaszöveg: Lev Tolsztoj: Anna Karenyina. Első rész, XII. 3-4. Officina Textologica 1998/2, 185–193.

Boda I. Károly, B. Porkoláb Judit (1998): Koreferenciális kifejezések és koreferenciarelációk. [Co-referential phrases and co-reference relations.] Példaszöveg: Szent János Apostol Jelenéseinek könyve. 21:9-23. Az új Jeruzsálem. (Részlet) Officina Textologica 1998/2, 32–56.

Petőfi S. János (1997): Egy poliglott szövegtani-szövegnyelvészeti kutatóprogram. [A polyglot textological / text linguistic research program.] Officina Textologica 1997/1. Debrecen: Kossuth Egyetemi Kiadó.

Petőfi S. János (1998): Koreferenciális elemek és koreferenciarelációk. [Co-referential elements and co-reference relations.] Példaszöveg: Mt. 9,9-13. Máté meghívása. Officina Textologica 1998/2, 15–31.

Petőfi S. János, Dobi Edit (1998): Utószó. [Epilogue.] Officina Textologica 1998/2, 238–261.

Tolcsvai Nagy Gábor (1998): Explicitség és koreferencia. [Expliciteness and co-reference.] Példaszöveg: egy anekdotikus prózarészlet. Officina Textologica 1998/2, 223–237.

Tuba Márta (1998): Koreferenciális kifejezések és koreferenciarelációk egy mesében. [Co-referential phrases and co-reference relations in a folk tale.] Példaszöveg: a kakas aranygarasa. (Részlet) Officina Textologica 1998/2, 142–168.